TECHNISCHE UNIVERSITÄT WIEN

Diplomarbeit

Reference Water Mapping based on Sentinel-1 and -2 data

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Geodäsie und Geoinformation

eingereicht von

Florian Roth BSc
Matrikelnummer 01126275

ausgeführt am Department für Geodäsie und Geoinformation
der Fakultät für Mathematik und Geoinformation
der Technischen Universität Wien

Betreuung
Betreuer: Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Wagner
Mitwirkung: Michaela Rättich M.Sc.

Wien, am 2. Januar 2020

(Unterschrift VerfasserIn) (Unterschrift BetreuerIn)

Acknowledgements

The writing of this thesis would not have been possible without the support of a variety of people and I want to express my gratitude to them.

At first I want to thank my supervisor Professor Wolfgang Wagner from TU Wien for his support and for enabling my stay abroad in Germany. Additionally I would like to thank the research group 'Natural Hazards' from the German Aerospace Centre (DLR). Especially I want to mention Michaela Bettinger, Sandro Martinis, Stefan Schlaffer and Marc Wieland for their professional support even after I left the DLR.

Special thanks go to my friends and family. My stay in Germany would not have been possible without the support from my girlfriend Stefanie. I am very thankful for sharing all joys and difficulties of the everyday student life with you. I am looking forward to our further joint path.

My parents helped me to develop as a person and they have supported me through my whole studies. Both of you and my sister Veronika deserve my huge gratitude! I would also like to thank Stefanie's family who supported me like I had always been part of their family.

I would like to mention my friends since school Alexander, Michael, Reinhold and Roland. Although some of us went to different universities and some do not live in Vienna anymore you all have a part in this thesis. Finally I express my gratitude to my new friends, who I got to know at DLR. In particular I would like to thank Gabriel, Henri, Nicole and Stefan for their support.

Abstract

Flood events are one of the most dangerous natural hazards on Earth [32]. By using instruments of Earth Observation, information about the actual water extent and the event's process can be provided. Subsequently, a targeted rescue of victims can be supported and the potential damages can be reduced. An important task is to determine the event's true surface water extent and to distinguish between areas which are affected by an event and those which are generally inundated. Therefore observations obtained during a flood event are compared to reference information which represents the open surface water bodies under normal conditions. Uncertainties within the reference information could directly distort the result of flood detection. In this thesis multiple methods for defining valuable reference information are observed and compared. One focus lies on the diverging properties between surface water detection based on optical (Sentinel-2) or radar observations (Sentinel-1). Therefore Sentinel-1 and Sentinel-2 water maps were processed by fully automatic processing chains developed by the German Aerospace Centre's (DLR) research group 'Natural Hazards'. A flexible work-flow for defining reference information from time-series of water maps is presented. As a result diverging properties of reference information calculated from different sensors, time periods and other adaptions of the work-flow are presented. By the use of these findings the thesis should support potential use of satellite based reference water mapping. The aim is to improve the understanding of characteristics of certain parameters like the used sensors or the considered period of time. Finally, an example for the use of sensor specific differences in flood detection is given. All results are tested in Bihar (India) where temporal changes of the surface water areas appear quite often. The second test region is the region surrounding Zaragoza (Spain) where the water areas are hardly affected by temporal changes.

Zusammenfassung

Hochwasserereignisse stellen mitunter die gefährlichsten Naturgefahren für das menschliche Leben dar. Methoden der Erdbeobachtung können Daten liefern, um den tatsächlichen Wasserstand zu bestimmen und den Fortschritt des Ereignisses beobachten zu können. Dadurch ist eine gezieltere Rettung der Betroffenen möglich und zusätzlicher Schaden kann reduziert werden. Eine entscheidende Aufgabe dabei ist die Definition der tatsächlichen Wasserflächen und im Weiteren die Unterschiedung in jene Bereiche, die unter normalen Umständen von Wasser bedeckt sind und jenen, die nur während der Flut überschwemmt sind. Zu diesem Zweck wird die erfasste Wasserfläche zum Zeitpunkt des Hochwasserereignis mit einer Referenzfläche verglichen, welche das Gebiet unter normalen hydrologischen Bedingungen repräsentiert. Ungenauigkeiten in dieser Referenzfläche könnten das Ergebnis verfälschen. In dieser Arbeit sollen unter schiedliche Methoden zur Berechnung einer Referenzwassermaske verglichen werden. Besonderes Augenmerk wird dabei auf den Einfluss von optischen bzw. radarbasierten Beobachtungen gelegt.

Dafür wurden Sentinel-1 und Sentinel-2 Wassermasken mithilfe der automatischen Prozessoren berechnet, welche von der Forschungsgruppe "Naturgefahren" des Deutschen Zentrums für Luft- und Raumfahrt (DLR) entwickelt wurden. Auf dieser Grundlage wurde ein flexibler Arbeitsablauf definiert, der Referenzwassermasken aus entsprechenden Zeitserien bestimmt.

Unterschiede in resultierenden Referenzwassermasken, welche mit verschiedenen Sensoren, zeitlichen Abschnitten und anderen Parametern bestimmt wurden, können als Ergebnis dieser Arbeit betrachtet werden. Mit diesen Erkenntnissen soll das Verständnis verschiedener Einflussfaktoren auf die Definition einer Referenzwassermaske verbessert werden. Zum Abschluss wird eine Möglichkeit des Einsatzes der sensortypischen Eigenschaften anhand eines Hochwasserereignisses überprüft. Als Interessensgebiete wurden Bihar (Indien) und Zaragoza (Spanien) gewählt, wobei die Wasserflächen in Bihar starken zeitlichen Änderungen unterworfen sind, während Zaragozas Wasserflächen zeitlich konstanter sind.

Contents

List of Acronyms

Acronyms

CCI Climate Change Initiative. 17, 19

CEMS Copernicus Emergency Management Service. 15, 18, 60

CNN Convolutional Neural Network. 23, 24

Copernicus WB Copernicus Global Land Service Collection 2. 17, 19

DLR German Aerospace Center. 8, 21–23, 29

EFAS European Flood Awareness System. 18

EO Earth Observation. 8, 11, 58

ESA European Space Agency. 13, 17, 19, 22, 23

G3WBM Global 3 arc-second Water Body Map. 16, 19, 28

GDAL Geospatial Data Abstraction Library. 26, 27, 29

GFC Global Forest Change. 16, 17, 19

GIEMS Global Inundation Extent from Multi-Satellites. 8, 17–19, 60

GIW Global Inland Water. 16, 17, 19

GLOFAS Global Flood Awareness System. 18

GLOWABO Global Water Bodies Database. 16, 17, 19

GLWD Global Lakes and Wetlands Database. 17

GSWE Global Surface Water Explorer. 16–21, 28, 37, 47, 51–53, 60, 61

HAND Height Above Nearest Drainage. 22, 23

MMAP Months of median average precipitation. 34, 44, 45

MMWC Months of median water cover. 34, 44, 45, 51

MNDWI Modified Normalised Difference Water Index. 11

MOD44W MODIS land-water mask. 17, 19, 22

MODIS Moderate Resolution Imaging Spectrometer. 12, 17–19

NDVI Normalised Difference Vegetation Index. 11

NDWI Normalised Difference Water Index. 11, 12

S-1 Sentinel-1. 8, 9, 12, 13, 19, 21–25, 29, 30, 34, 35, 37–43, 46, 47, 49–61

S-2 Sentinel-2. 8, 9, 12, 19–25, 29, 30, 34, 35, 37–41, 43, 44, 46–53, 55, 56, 58–61

SAR Synthetic Aperture Radar. 8, 9, 13, 14, 16, 17, 23, 34, 40, 54, 61

SAR-WBI SAR-based Water Body Indicator. 16, 17, 19

SRTM Shuttle Radar Topography Mission. 16, 19, 22, 23

SWBD SRTM Water Body Dataset. 16, 17, 19, 22, 23, 29, 37, 51–53, 60

SWOT Surface Water and Ocean Topography mission. 13

WGS84 World Geodetic System 1984. 23, 29

WO Water Occurrence. 30–32, 35, 41, 43, 48, 49, 58–60

1 Introduction

1.1 Motivation

Many regions of the world are at risk of flood events and many people are directly or indirectly affected by them. The United Nations stated that more people lose their lives due to floods than due to any other natural hazard [32]. Additionally, massive destruction of infrastructure has to be expected during a hydrological extreme event [32]. In Figure 1 a global comparison of the occurrence of different natural hazards is given. It can be seen that flood events increased in frequency during the last decades and present now the most frequent disaster event. Earth Observation (EO) can help to provide important information to authorities or non-governmental organisations to coordinate emergency management activities and to monitor the event's process as well as to reduce potential damages. The extend of inundation is one of the key variables in flood disaster management and can be observed using a variety of sensors or methods [43]. To differentiate between areas, which are only temporal limited inundated during a flood event, and areas, which are inundated under normal conditions, a reference water map is needed. There are multiple global and local datasets available for this purpose, varying in spatial resolution, accuracy, update frequency and many other measures. Especially in regions, where temporal changes of surface water are common (= dynamic surface water) a reference dataset with a high spatial and temporal resolution is needed for an appropriate flood analysis. Consequently, the reference map has to be unaffected by hydrological extreme events but it should consider the seasonal variability of the surface water.

Due to the social and scientific significance of fresh water many disciplines have interest in the spatial distribution and extent of surface water bodies [33]. To satisfy multiple conditions of expert users (including hydrologists, water and disaster managers or climate scientists) a variety of methods and datasets have been developed in recent years [19]. Beside optical observations, Synthetic Aperture Radar (SAR) is an often used source to retrieve surface water bodies [19]. Based on the used sensor the resulting water maps suffer from different limitations. While optical observations are affected by cloud cover and limitations from vegetated areas [19], SAR observations have to be refined locally [42]. To improve the interpretation of inundated areas during a flood detection process based on remote sensing those sensor specific weaknesses have to be known. The Global Inundation Extent from Multi-Satellites (GIEMS) dataset combines multiple radar sensors with optical observations to correct those weaknesses and to avoid local refinements [39]. Taking up this approach a method to benefit from Sentinel-1 (S-1) and Sentinel-2 (S-2) information is presented within this thesis.

The research group 'Natural Hazards' in the department 'Geo-Risks and Civil Security' of the German Aerospace Center (DLR) [4] produces crisis-relevant information products based on Earth observation data in the case of natural disasters, e.g. floods, wildfires, volcanic eruptions, landslides, etc. Therefore, the group developed fully automatic processing chains detecting surface water bodies from satellite data. Two of them are of particular importance for the derivation of the reference water map due to their

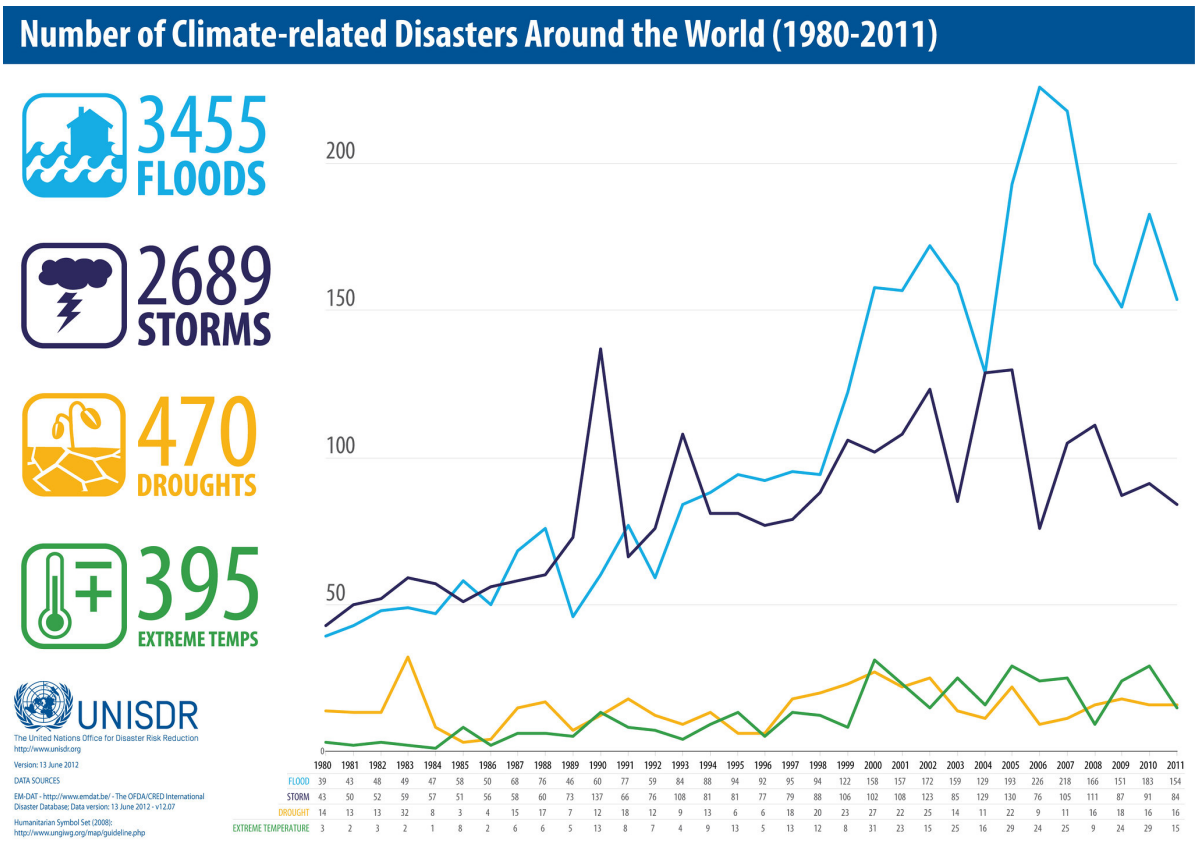


Figure 1: Global disaster events (1980 - 2011) [14]

high spatial and temporal resolution: One of them uses C-band SAR data observed by S-1 [44] while the other takes multi-spectral data from S-2 as input for the derivation of the flood extent [47].

1.2 Objectives

This thesis is focused on supporting the detection and interpretation of inundated areas during flood events which is one of the main research topics of the research group 'Natural Hazards' [4]. A work-flow for calculating reference information, which represents the observed surface water bodies under normal conditions, is presented. Therefore, S-1 and S-2-based water maps, produced by the group's processing chains, can be combined or considered individually. Additionally, possible parameters for the variation of the work-flow and their effects on the final result are shown. Finally, an identification of parameter and sensor specific differences in reference water mapping and the corresponding influences to disaster management and long-time analysis of flood events will be investigated. Due to the absence of ground-truth data of the water extent under normal conditions and numerous definitions of what is considered as "good reference" [29] research misses the possibility to quantify the quality of reference water information. Consequently, the detected parameter and sensor specific differences of this thesis aim

to support potential users of flood detection processes. This approach should enable the user to interpret reference water map's reliability based on the used parameter or sensor and to detect potential weaknesses. The main objectives of this thesis are summarised below:

- Developing a flexible work-flow for the calculation of reference water information from given water maps and enable the combination of multiple sensor systems.
- Presenting possible variations of the developed work-flow including an automatic threshold definition.
- Determine method or sensor specific differences in the resulting reference water maps which can be used to interpret the reliability of a flood detection result.

1.3 Structure

An overview of commonly used methods to detect water areas by the use of satellite data and available reference water products are described in Chapter 2. Used datasets, considered areas of interest and algorithmic tools are topic of Chapter 3. The work-flow, which was developed in the context of this thesis, is presented in Chapter 4. The comparison of multiple variations of the presented workflow and the corresponding interpretation is given in Chapter 5 as well as the use of the findings within a flood detection process. A summary of the thesis' findings and an outlook are part of the conclusion in Chapter 6.

2 State of the art in satellite-based reference water mapping

In this section state-of-the-art methods for surface water detection and the definition of a valid reference information is presented. Finally available surface water datasets and flood detection services are introduced and discussed.

2.1 Principles of surface water detection

Although earth's inland surface water areas represent only 4 % of earth's surface, numerous use cases for the corresponding EO data exist [42]. Due to diverging needs of multiple potential users (e.g.: hydrologists, water and disaster managers or climate scientists [19]) of surface water extent and distribution there exist a large number of water detection approaches. Spatial and temporal resolution have to be chosen depending on the specific use case as well as the corresponding satellite sensor. In general optical and microwave sensors can be used for surface water detection or flood mapping [43].

2.1.1 Water detection by the use of optical sensors

Optical satellite observations are the most common way for detecting surface water bodies [19]. Literature is dominated by two traditional approaches to extract open water bodies from optical data: Threshold and supervised classification [40]. Using thresholds for single spectral bands (e.g.: near-infrared) or remote sensing derived indexes serve as instruments to extract water areas in the first approach [40]. Traditionally, indexes like the Normalised Difference Water Index (NDWI), the Normalised Difference Vegetation Index (NDVI) and the Modified Normalised Difference Water Index (MNDWI) well distinguish between water and non-water areas [40]. An example for the NDWI can be seen in Figure 2b, while the corresponding RGB images is given in Figure 2c. Equations 1 [49], 2 [49] and 3 [50] show how these indexes can be calculated:

$$NDWI = \frac{\rho_{\mathrm{G}} - \rho_{\mathrm{NIR}}}{\rho_{\mathrm{G}} + \rho_{\mathrm{NIR}}}, \tag{1}$$

$$MNDWI = \frac{\rho_{\mathrm{G}} - \rho_{\mathrm{MIR}}}{\rho_{\mathrm{G}} + \rho_{\mathrm{MIR}}}, \tag{2}$$

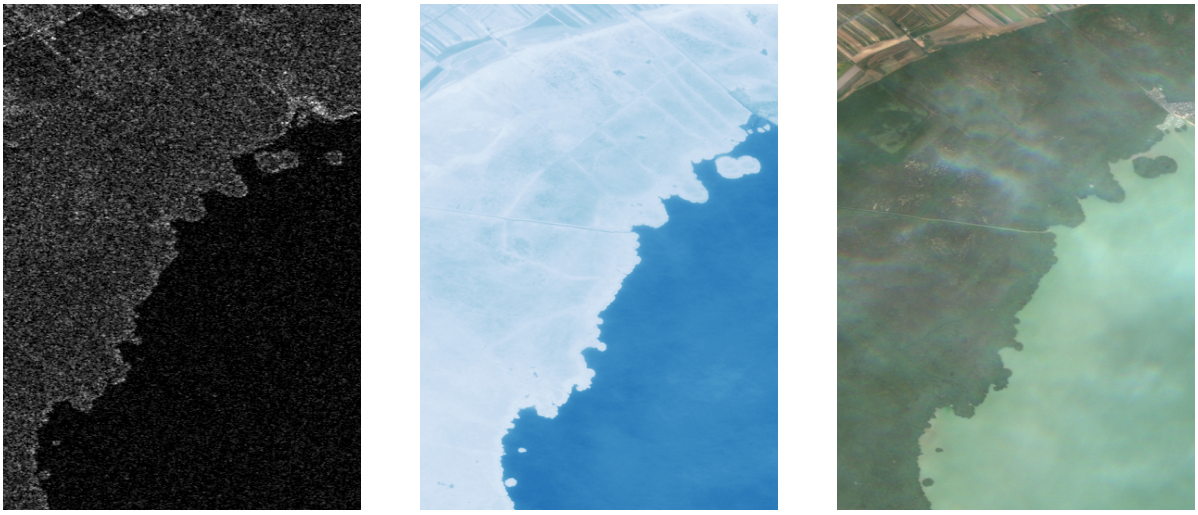
$$NDVI = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{R}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{R}}}, \tag{3}$$

where ρ_{G} is the green spectral band, ρ_{MIR} is the middle infrared band, ρ_{NIR} is the near infrared and ρ_{R} is the red band.

The second approach of extracting water bodies from optical observation is to perform a supervised classification [40]. The idea behind supervised classification is pattern recognition or machine learning. A classifier is trained by analysing features of already defined land-cover classes (e.g. water areas). The considered features can be spectral bands or indexes, similar to the first approach. Once the classifier is trained adequately

it can be used to perform classification tasks (e.g. extract water areas) [22]. Commonly used supervised classification methods are: Support Vector Machines (SVM), Maximum Likelihood (ML), Decision Tree or Random Forest (RF) [40].

Based on the specific use case a variety of satellite based optical sensors is available and can be used for water detection purposes. Regarding special use cases like climate modelling a high temporal resolution is of high interest to observe the underlying seasonal dynamics of water bodies [43]. Although the Moderate Resolution Imaging Spectrometer (MODIS) provides a coarse spatial resolution, it is particular suitable for this purpose due to its global coverage every two days [19]. Since the first Landsat satellite was launched in 1972, Landsat images have become the primary source for change detection and long-term analysis applications as well as medium spatial resolution (30 m) multi-spectral imagery. Due to systematic global coverage at high revisit rate of 5 days and a spatial resolution of up to 10 meters S-2 is of high potential for global water detection systems. It delivers continuously multi-spectral optical observations in 13 spectral bands and is already used in many modern applications (e.g.: [47]) [25].



(a) S-1 backscatter image (darker pixels represent a lower backscatter)

(b) S-2 NDWI (blue represents a higher NDWI)

(c) S-2 RGB image

Figure 2: Coast of Lake Neusiedl in Austria observed by radar (S-1 observed at 2019-8-28) and optical (S-2 observed at 2019-8-24) sensors (loaded from Copernicus Open Access Hub [2])

2.1.2 Water detection by the use of radar sensors

As described in [43], radar radiation is reflected away from the sensor when hitting an open water surface (= specular reflection). Due to this effect, a low backscatter value is measured which can be used as distinguishing feature for water detection (example is given in Figure 2a). For water detection a suitable threshold is defined to extract regions

of lower backscatter observations. This value can either be found by visual inspection or an automatic algorithm [43]. Due to the backscatter's fluctuation caused by the speckle effect [1] or environmental impact (e.g.: wind [44]) SAR-based flood detection is often progressed by change detection approaches [29].

Microwave remote sensing distinguishes between active and passive radar systems [48]. While an active system transmits pulses of microwave radiation and measures the reflected signal, a passive system measures the microwave energy radiated or reflected naturally by the earth's surface [48]. Besides of active observations (as described above), passive observations can be used for water detection as well. In [39], passive radar observations are used as primary source for water detection and active radar observations as well as visible and near infrared measurements are added for improvement reasons [19].

ENVISAT Advanced SAR (ASAR) sensor was widely used for applications using moderate-resolution radar images until the mission's end in April 2012 [43]. Since the launch of the first S-1 satellite in April 2014 continuity of global radar images is ensured by the European Space Agency (ESA) satellite [43]. ENVISAT Advanced SAR (ASAR) and S-1 both operate at C-band frequency [2]. Higher spatial resolution can be obtained from TerraSAR-X or Cosmos-SkyMed but free data policy of ESA is beneficial especially for global or multi-temporal needs. NASA's Surface Water and Ocean Topography mission (SWOT) will be launched in 2020 and is designed specifically to improve surface water detection by using K-band radar [3] [18].

2.1.3 Differences and potential synergies of optical and radar sensors

Although optical and SAR data can both be used for water detection purposes the resulting water extent varies due to diverging characteristics of the sensor types. While SAR is able to penetrate clouds, heavy rainfall and vegetation, optical data suffers from the absence of information within those areas [19]. It has to be mentioned that radar observations are also slightly effected by environmental effects like vegetation or strong winds which can increase the backscatter value locally [43]. Winds could increase surface roughness of the water surface which prevents specular reflection [43]. Potential increase of backscatter caused by vegetation or infrastructure could be produced by double bounce effects [43]. Radar measurements can operate day and night because they are not dependent on sunlight [19]. On the other hand classification algorithms, which are using optical data, have to deal with variable spectral signatures across different types of surface water [42]. Due to lower environmental sensitivity, radar measurements are of significant importance during extreme conditions like heavy rainfall in subtropical rainy season [19], volcanic eruptions or shallow solar illumination angles in high-latitude regions [42]. Because of the fact that floods are often driven by heavy rainfalls, SAR

[1] Radar images are affected by the speckle effect which looks like salt and pepper noise. It can be found in all coherent imaging systems and is caused by the interference of multiple scatterer's echos within a resolution cell [48].

[2] C-band is located at a wavelength interval of 5.22 cm to 7.14 cm [31]

[3] K-band is located at a wavelength interval of 8.3 mm to 2.75 cm [31]

observations are sometimes the remaining data source for satellite-based analysis [43]. There tends to be a lower number of radar based global applications due to the need of local calibrations to improve accuracy and the lack of full global data availability in the past [19]. Many global water detection approaches based on SAR must deal with systematic commission errors caused by land cover types which showed similar backscatter behaviour as water surfaces [42]. In general, it is difficult to distinguish very saturated soil from standing open water bodies based on radar observations [19]. Additionally ice caps, arid zones, salt pans and deserts (in particular sandy deserts) could be confused with water covered regions [42]. A summary of mentioned sensor type's advantages in the field of satellite-based water detection is given in table 1 and 2.

Advantages	**Disadvantages**
There is less local calibration necessary to detect surface water globally.	There is an absence of sub-canopy information and cloud penetration ability.
There are long time-series of optical observations in medium spatial resolution available (e.g.: Landsat).	The retrieval of information requires the presence of sunlight.
Less probability for confusion with other land cover classes (e.g.: ice caps or deserts) or saturated soil.	Less suitable during extreme weather events.

Table 1: Summary of advantages and disadvantages of optical data for water detection

Advantages	**Disadvantages**
Detection of open water underneath clouds or vegetation is possible.	Local variations of the backscatter coefficient due to wind, vegetation or urban infrastructure.
The detection is not dependent on sunlight or surface water type.	Local calibration is needed to provide global coverage.
Data can be provided during extreme weather events.	Possible confusion of several land cover classes and open water bodies.

Table 2: Summary of advantages and disadvantages of SAR data for water detection

One can see that optical and SAR based water detection show different characteristics and advantages. Based on this knowledge the sensor which fits best for a specific scenario (area of interest and given time interval) can be chosen. Additionally, sensor specific characteristics can be used to improve the general result as it was done in [33] or [39]. In this thesis the knowledge of the differences of radar and optical observations is used to improve flood detection.

2.2 Methods for the generation of reference water information

'Flood' is defined as temporary covering of land by water which is not normally covered by water [7]. As mentioned in [44], knowledge of flood extent is of significant importance for effective disaster management, helping humanitarian relief organisations and decision-makers. Additionally these data can be used in various other disciplines like hydraulic modelling, flood prevention activities, insurance risk management, spatial planning [44] or climate modelling [43].

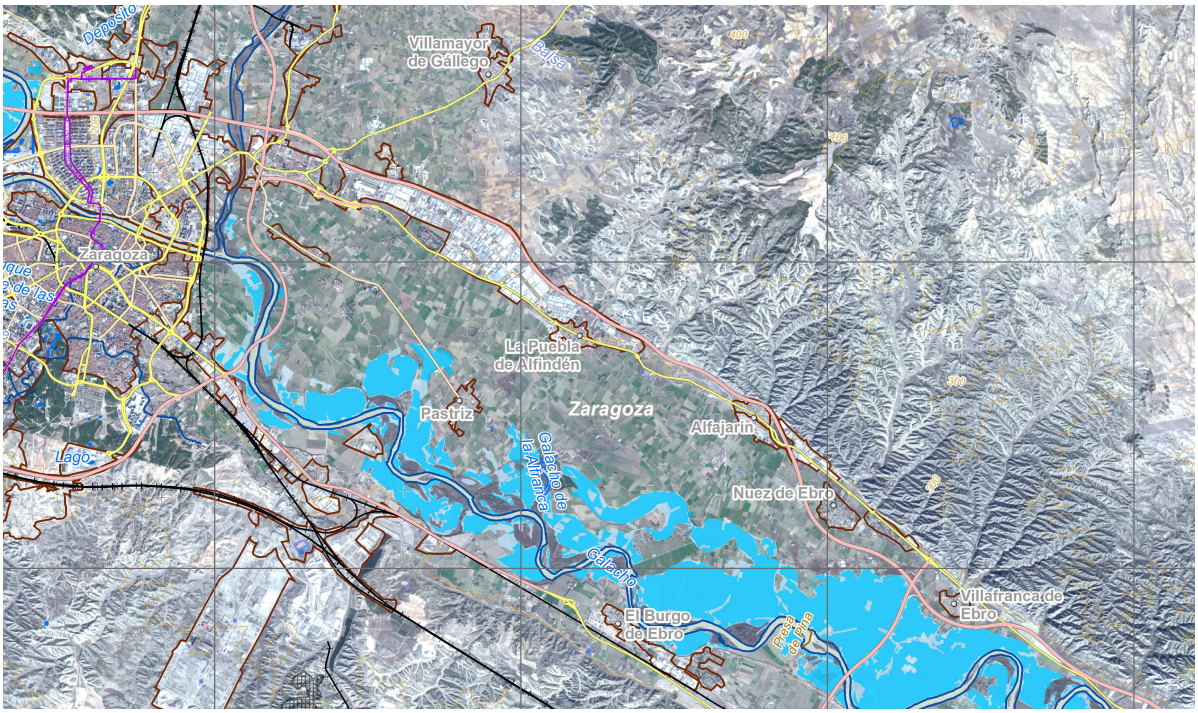


Figure 3: Flood Mapping of Zaragoza (Copernicus Emergency Management Service (CEMS) at 2018-4-16 [5])

According to [29], flood extent is observed by the comparison of the water extent during a flood event and the affected region under normal conditions. In Figure 3 an example of detected flooded areas conducted by the CEMS [5] is given. The blue regions represent the flooded areas which are obtained from subtraction of normal water condition areas. In literature numerous definitions of normal conditions can be found due to the fact that a "good reference" is not commonly defined. Frequently flood extent is calculated by taking the difference between two water maps, using pre-flooded images as reference. Unfortunately in the past only little effort was put into finding the most adequate reference image, although a high risk of over- or underestimation of the true flood extend is given [29]. If the water extent of the selected reference image differ from normal conditions, flood detection would be affected accordingly. In [29], the negative impacts on flood detection are shown when manually selecting the corresponding reference image. Often the first non-flooded image is selected by visual inspection although it is not necessarily the most adequate one [29]. An improvement can be achieved by

the use of long-term satellite data archives [29]. Some approaches automatically select the historical driest image from the archive as reference (for the example of soil moisture given in [37]). To increase cost-efficiency in case of high resolution imagery, statistical approaches are applied from [29] and [36]. In [29], two statistical indexes based on radar backscatter measurements are calculated. Further on a successful sensitivity analysis is done to figure out whether the statistical approach yields a higher flood classification accuracy. In [36], a SAR reference image is selected by calculating the Jensen-Shannon (JS) divergence-based index.

An alternative to the already mentioned difference method is the multi-temporal modelling of normal conditions at the area of interest by analysing a corresponding time-series. In [43], a harmonic analysis of a SAR time-series is done and in [45] a multi-temporal approach from optical data was performed. Another concept is the usage of existing global surface water databases like SRTM Water Body Dataset (SWBD) [13], Global Surface Water Explorer (GSWE) [38] or Global 3 arc-second Water Body Map (G3WBM) [50] as reference. A discussion of a selection of these datasets is provided in the following section (2.3).

2.3 Global surface water data for reference water mapping

During the last decade, a large number of global or near-global datasets of open water bodies have been developed and published. They can be organised by spatial resolution, extent, considered time period, sensor type used, types of secondary products (like water occurrence or maximum water extent) and others. The first near-global dataset and still one of the most used surface water products is SWBD [13]. Active radar observations from the Shuttle Radar Topography Mission (SRTM) were used to generate this map (SWBD) of surface water bodies between a latitude of 60°N to 54°S at 90 m spatial resolution [33]. SWBD was observed between 11th and 21st February 2000. As mentioned in [23] and [42], modern water detection approaches revealed the potential to update some areas within SWBD.

Another SAR-based dataset is SAR-based Water Body Indicator (SAR-WBI) which uses Envisat ASAR data from 2005 to 2012 and provides permanent open water bodies for the corresponding time interval. SAR-WBI covers latitudes between 84°N and 60°S and shows a spatial resolution of 150m [33].

Considering different periods of time, multiple Landsat-based products are available, for example: Global Forest Change (GFC) [28] (period 2000-2012), Global Inland Water (GIW) [26] (2000 epoch), G3WBM [50] (1990-2010), Global Water Bodies Database (GLOWABO) [46] (2000 epoch) or GSWE [38] (1984-2015 or 1984-2018) which are providing a spatial resolution of 14.24 - 90m [33]. Recently, the European Commission's Joint Research Centre released a new version of GSWE which covers a period from 1984 to 2018.

In the scope of this thesis, G3WBM and GSWE must be highlighted because they display the parameter *water occurrence* which is used as basis for calculating reference water maps (section 4.1.4). Both datasets overtop all other mentioned datasets with respect to their considered time period's length. Although both datasets provide similar

results the remaining differences can be explained by different periods of time used to calculate the datasets [19]. GSWE presented a monthly normalisation method to balance the influence of the number of valid observations each month [19].

A different approach was applied by the GIEMS [39] by combining passive and active microwave data as well as visible and near infrared measurements [19]. The combination of multiple sensors helps to minimise the limitations and uncertainties of the individual sensors [19]. As mentioned in section 2.1.3, SAR based water detection tends to overestimate the water extent because of potential confusion with other land-cover classes. In fact, GIEMS shows less overestimation in regions of saturated soil but still includes sub-canopy water areas [19]. Originally GIEMS was published providing a 25km spatial resolution, but by applying downscaling methods the resolution was reduced to 500m (GIEMS-D15 [27]) and later 90m (GIEMS-D3 [18]) [19]. A global comparison of GSWE and GIEMS-D3 is shown in Figure 4. Positive values are coloured in red and represent the percentage of water areas that are detected only by GSWE while blue areas represent the percentage of water areas only detected by GIEMS-D3. In [19], local analysis in the Ganges-Brahmaputra region (highlighted area (1)) suggest that GIEMS show overestimation there which can be seen in Figure 4. An underestimation of the Landsat-based GSWE dataset, which is caused by vegetation, can be revealed in the region of the Amazon river (highlighted area (2)). Another comparison of surface water products is given by [33] which was further used to combine multiple datasets and to build an accurate mask of inland/ocean water for the 2000-2012 period in the framework of ESA's Climate Change Initiative (CCI). The selected datasets consist of SWBD, SAR-WBI, GFC and GIW [33].

A different sensor was used by MODIS land-water mask (MOD44W) which provides water body maps of 250m spatial resolution obtained by MODIS and SWBD [33]. Copernicus Global Land Service Collection 2 (Copernicus WB) processed PROBA-V observations to get surface water bodies with a spatial resolution of 1 km [33]. For the production of the Global Lakes and Wetlands Database (GLWD), large-scale and regional sources before 2000 were combined [33]. An overview of the described method's parameters is shown in Table 3.

One can notice that the datasets of this selection (Table 3) were produced with various reasons in mind. Depending on the user's intention a specific dataset might be suitable or not. Although GLOWABO offers the the best spatial resolution of this selection it might not be useful in case of hydrological dynamic regions because of the single year of considered time period. Although SWBD is affected by the same issue it still can be used for many purposes. For example it is used in [33] to replace water in correspondence to islands which are missing in the GFC dataset. Regarding applications which are focused on the dynamic behaviour of surface water areas the GSWE dataset would be a great option due to the long time period which was considered. Additionally GSWE provides useful products which are describing intra- and inter year changes of water areas [38]. It is noticeable that the only true-global datasets (latitudes from 90°N to 90°) of this selection are produced by combining multiple sensors. This might be caused by the diverging characteristics of the specific sensors (described in Section 2.1).

Besides of static water body maps, which can be used as reference information for

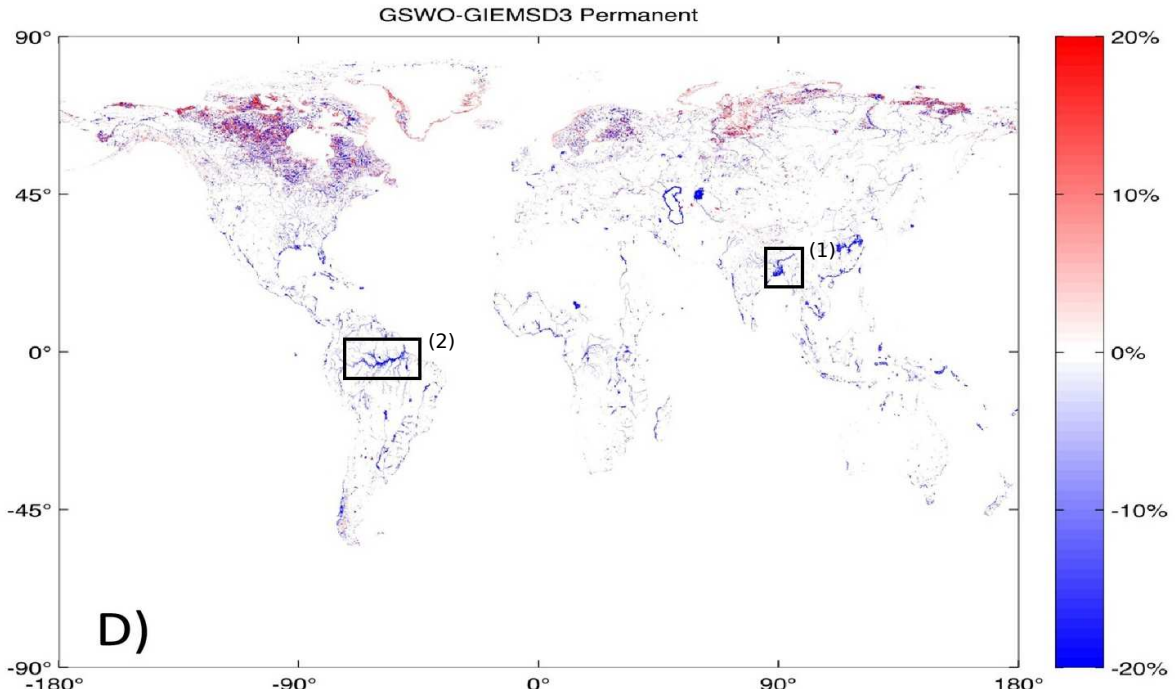


Figure 4: Global comparison of GSWE and GIEMS-D3 (in %) [19]

flood detection, flood services provide event-based flood extent data. The Global Flood Awareness System (GLOFAS) [9] and European Flood Awareness System (EFAS) [6] provide daily flood forecast information based on hydrological models. Both systems are available as CEMS. Dartmouth Flood Observatory on the other hand performs an event-based flood observation by the use of MODIS and other optical and radar sensors [3].

Dataset	Sensor/ Plattform	Latitudes	Spatial resolution	Considered time period
SWBD	SRTM	60°N to 54°S	90 m	February 2000
SAR-WBI	Envisat ASAR	84°N and 60°S	150 m	2005-2012
GFC	Landsat	80°N and 57°S	30 m	2000-2012
GIW	Landsat	90°N and 60°S	30 m	2000
G3WBM	Landsat	90°N and 60°S	90 m	1990-2010
GLOWABO	Landsat	60°N and 56°S	14.25 m (pan-sharpened)	2000
GSWE	Landsat	78°N and 60°S	30 m	1984-2015
GIEMS	multiple	90°N and 90°S	25 km - 90 m	
ESA CCI	multiple	90°N and 90°S	150 m	2000-2012
MOD44W	MODIS	60°N and 60°S	250 m	2000-2001
Copernicus WB	PROBA-V	80°N and 60°S	1 km	January 2014

Table 3: Overview of all discussed global surface water products and flood services

3 Materials and Datatsets

In this Chapter the areas of interest of this thesis are presented and a description of their hydrological behaviour is given (Section 3.1). The S-1 and S-2 based processing chains and their resulting water maps are introduced in Section 3.2. Finally, tools which are used to create the algorithm and the visualisation of the results are shown in Section 3.3.

3.1 Study areas

For the evaluation of multiple concepts to calculate reference water maps two areas of interest were selected: Bihar (India) and Zaragoza (Spain). Both areas differ in their climatic and hydrological situations. As a consequence, it can be explored which characteristics have to be considered when defining a reference water map in hydrological dynamic or more static regions.

3.1.1 Bihar

Bihar is an Indian state located in eastern India close to the border with Nepal. The state is divided by the Ganges River and its topology is dominated by the Gangetic Plain except of the foothills of the Himalaya in the northwest of Bihar. Multiple rivers flow from the Himalaya towards Ganges River whereby they frequently change channels and create high risk for destructive floods. Bihar is one of the most densely populated states in India which makes the high number of floods particularly dangerous for human lives. The subtropical region can be divided in three main weather seasons: hot weather season (March till mid-June), southwest monsoon season (mid-June till October) and

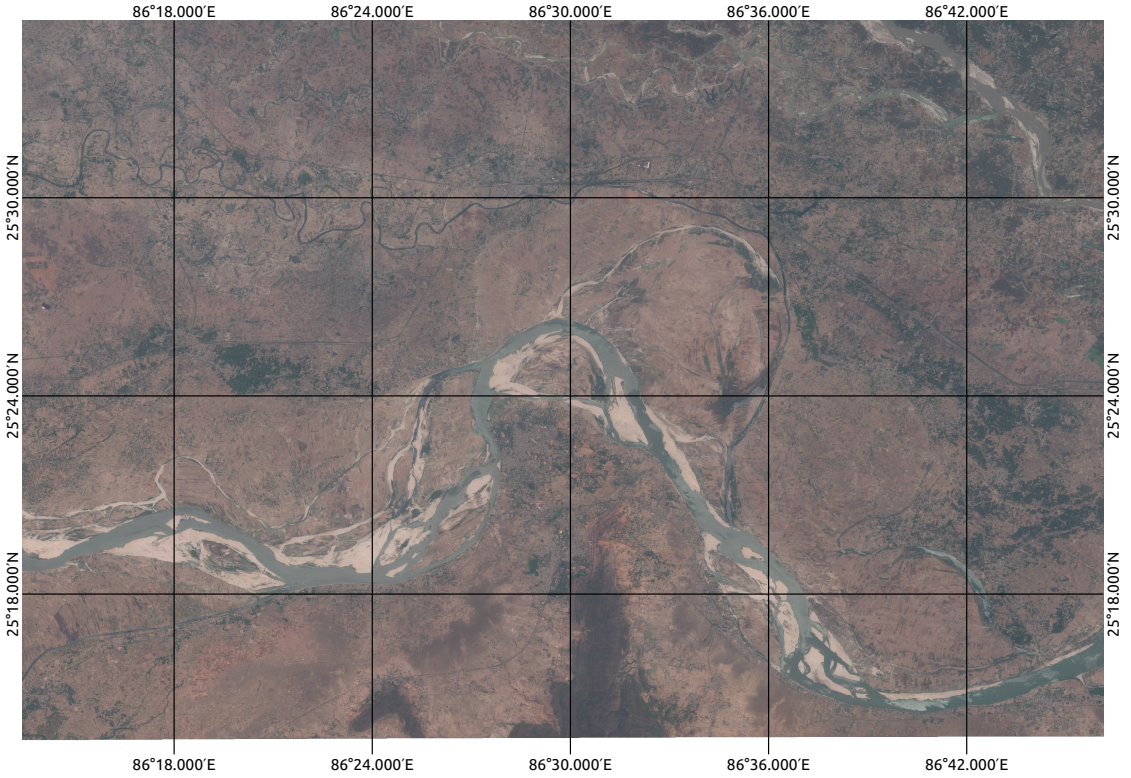
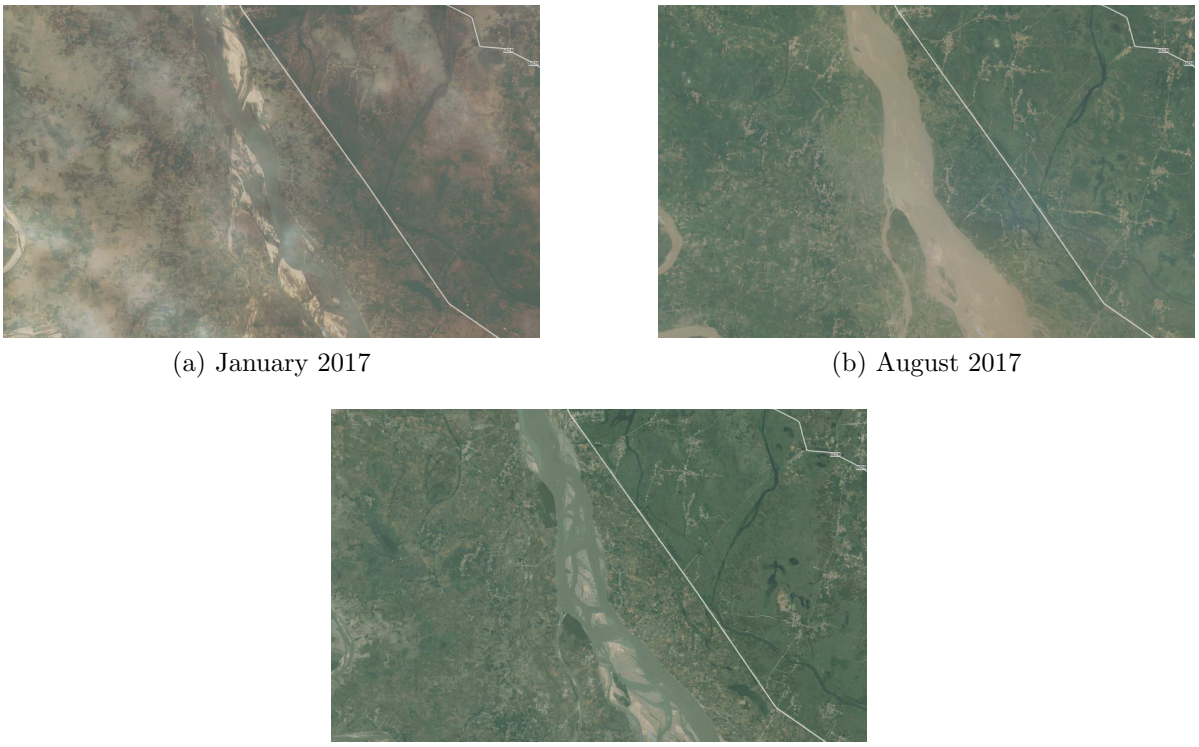


Figure 5: Overview of the first area of interest located in Bihar (S-2 true-colour image observed at 2019-5-18 and loaded from Open Sentinel Hub [2])

cold weather season (November till February). An overview of the specific area of interest is given in Figure 5 [24].

As described in Section 2.2, hydrological dynamic regions are particularly challenging in case of reference water mapping. As mentioned above, Bihar shows a dynamic hydrology and often suffers of flood events (variability is shown in Figure 6). The main challenge is the unequally distributed number of valid optical observations and the changes of river channels due to rainy season. The result could be incomplete or erroneous reference information which could cause an inaccurate determination of flood extent.

In case of Bihar years 2017 and 2018 were selected for further analysis because of the launch of a second S-2 satellite (Sentinel-2B) at March 7th 2017 [12]. By the use of this setup, the influence of an increased number of optical observation can be investigated. The year 2018 is also of interest because of the possibility to compare the results to the newly released GSWE dataset which covers the same year.



(a) January 2017

(b) August 2017

(c) November 2017

Figure 6: Variability of water extent and location in Bihar (images obtained from Planet [10])

3.1.2 Zaragoza

Zaragoza is a city in northeast of Spain and is the capital of the province of the same name [17]. The city is situated at the banks of Elbro River which is the longest river (910 km) and has the largest drainage basin of Spain [17]. In contrast to Bihar, the region around Zaragoza is hydrologically more static and it shows warm temperate climate. Elbro River is narrower than the Ganges River which allows to evaluate the impact of spatial resolution on the completeness of water bodies. Zaragoza was chosen to explore the results of reference water mapping in case of a hydrological static region and smaller river systems like they often occur in Europe. The year 2018 was selected to have an overlapping time period with the data obtained for Bihar. As already mentioned above (Section 3.1.1), the results can be compared to the newly released GSWE dataset from 2018. An overview of the specific area of interest is given in Figure 7.

3.2 Datasets

In this section all required input data for the algorithm which was developed for this thesis is presented (for the algorithm details see Chapter 4). Processed and provided by the DLR's research group "Natural hazards" [4], binary water maps from S-1 and S-2

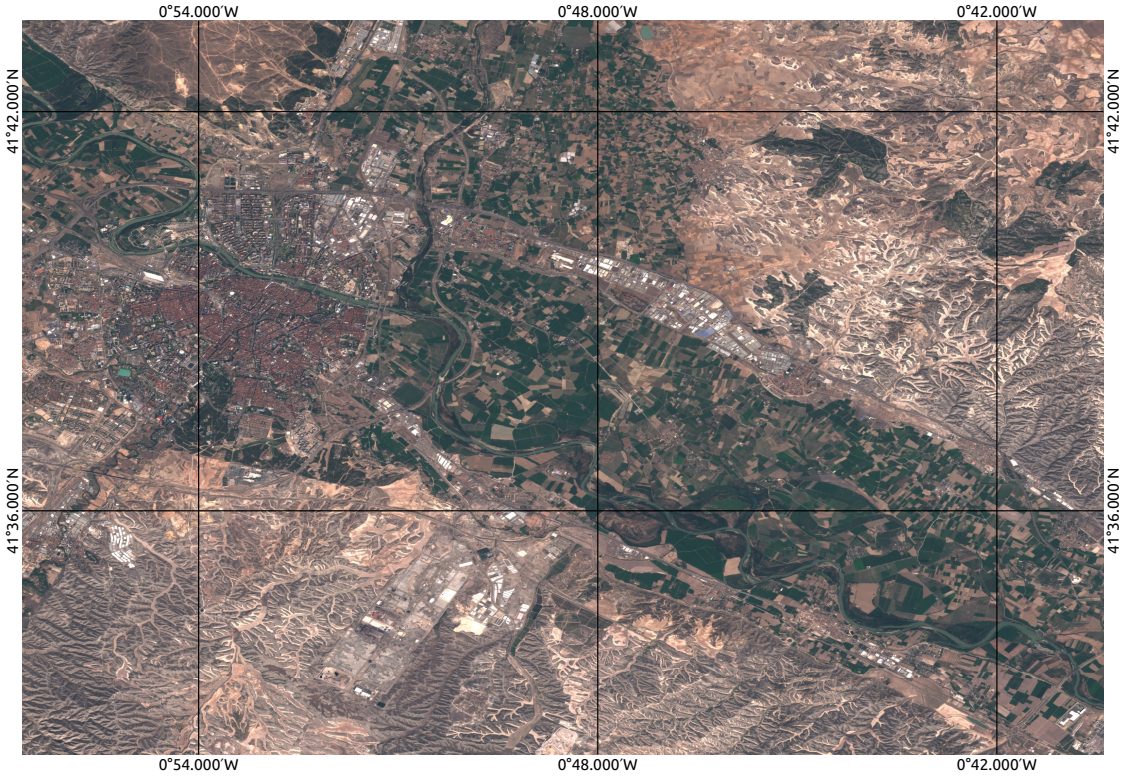


Figure 7: Overview of Zaragoza (S-2 true-colour image observed at 2019-9-6 and loaded from Open Sentinel Hub [2])

for both areas of interest (see Section 3.1) were used. Both datasets were produced by the corresponding fully automatic processing chains for detecting surface water bodies. In Section 3.2.1 the fully automatic processing chain based on S-1 [44] is shown, while section 3.2.2 presents the processing chain based on S-2 [47]. Additionally WorldClim's precipitation data, which is used in Section 4.2.1, is presented in Section 3.2.3.

3.2.1 Sentinel-1 water maps

To automatically retrieve surface water bodies from S-1 data the DLR's processing chain requires the following 6 calculation steps [44]:

Data ingestion Level 1 S-1 ground range detected (GRD) data, which are acquired in IW mode, are loaded automatically from ESA Sentinel Data Hub. The C-band radar data of S-1, which match user-defined criteria, loaded to the local file system and are prepared for further processing. Additionally, SRTM data, the Height Above Nearest Drainage (HAND) index [35] which is provided by the Hydroshed mapping product, SWBD data and MOD44W is needed.

Geometric correction and radiometric calibration Range-Doppler terrain correction [4] and radiometric calibration to sigma naught (dB) [5] are performed using the instruments from ESA's Sentinel Application Platform (SNAP). By the use of SNAP's Graph Processing Tool (GTP) this tasks can be automated.

Initial classification To divide the acquired dataset into land and water, initially an automatic tile-based thresholding procedure is performed. To reduce the risk to derive data from non-water areas the automatic threshold definition is supported by the HAND exclusion mask [44]. Terrain indices like HAND can be used to improve surface water mapping by removing water-lookalikes [30] [6].

Fuzzy-logic-based classification refinement Based on the backscatter level, relative elevation, slope information and size of an individual flood object the degree (real number within the interval of 0 to 1) of a membership to the water class is defined. The initial water bodies are classified by a threshold of the membership degree ($>$ 0.6). Additionally a region-growing step is performed.

Final classification A threshold ($\geq 15\,\mathrm{m}$) of the HAND index should reduce the amount of wrong classified elements within this step. Optionally a flood detection can be performed based on SWBD water bodies.

Dissemination of the results The processing chain's results are stored within a database and are published via Web Mapping Service (WMS) [15].

The results are binary images (=water maps) which distinguish water ('1') from land coverage ('0'). S-1 data provides a spatial sampling of 10 m but the SAR sensor originally resolves with $20\,\mathrm{m} \times 22\,\mathrm{m}$. As the processing is done in geographical coordinates, the processing chain yields a pixel size of a third arc seconds (approximately 30 m at the equator). Finally, the resulting maps are delivered in World Geodetic System 1984 (WGS84). An example of an input water map is given in Figure 8.

3.2.2 Sentinel-2 water maps

Beside water maps, which were processed from S-1 observations (described in Section 3.2.1), water maps based on S-2 observations are needed. Those maps were produced by DLR's S-2 automatic processing chain which was developed by the research group "Natural hazards" [47]. For segmentation purposes a U-net Convolutional Neural Network (CNN) [41] is trained based on 6 spectral bands (Red, Green, Blue, NIR, SWIR1

[4] Geometrical distortions of the Level 1 radar image has to be rectified by the use of digital elevation models (in this case SRTM) [48].

[5] Based on a digital elevation model (in this case SRTM) the actual incidence angle of the radar radiation is calculated. Further on the measurements are normalised relative to an area and transformed to decibel scale [48].

[6] The HAND index is widely used and recommended for large scale or global scale water mapping. It is defined by the elevation difference between a pixel an its associated nearest drainage pixel. In case of more complex terrain areas the Multi-resolution Valley Bottom Flatness (MrVBF) should be preferred [30].

and SWIR2) observed by S-2 [47]. Training data and hyperparameter were optimised specifically for water detection. Although U-net CNN was developed for bio-medical image segmentation it provides some advantages for the classification of remote sensing products [47]. Training and inference times can be minimised because only few data are needed for learning purposes [47]. Fortunately, the lightweight network is still able to achieve state-of-the-art results and meet the requirements of an emergency response application [47]. The spectral information, in the form of a high-dimensional feature vector, is taken as input for the U-net CNN [47]. The output of the network is a vector which contains the probability for each class and the resulting class is represented by the maximum probability [47]. The processing chain considers the following land-cover classes: cloud, cloud shadow, snow/ice, water and land [47].

Within this thesis the water class is exported to function as water maps. To distinguish between valid and invalid S-2 pixels a valid mask is processed for each scene individually. Therefore the classes *cloud*, *cloud shadow* and *ice/snow* are defined as invalid (=0) while *land* and *water* build the valid pixels (=1). Both datasets are exported as binary map in UTM projection. The spatial sampling is set to 30 m which matches Landsat's images. A water map is exemplary presented in Figure 8. Regardless of the used sensor (in this thesis S-1 and S-2) the resulting water maps appear as binary images and feature similar spatial sampling (30 m).

(a) Binary water map (white = water, black = non-water)

(b) Water map polygon (overlaying S-2 true-colour image from same date)

Figure 8: Exemplary reference water map result from Zaragoza (date: 2018-1-11)

3.2.3 WorldClim data

Beside water maps from S-1 (Section 3.2.1) and S-2 (Section 3.2.2), precipitation data is needed within this thesis. In Section 4.2.1, multiple approaches to automatically define useful time periods for reference water map processing are presented. To examine whether precipitation data can be used to find time periods which represent normal hydrological conditions WorldClim precipitation datasets were loaded.

WorldClim provides climate layers like temperature or precipitation at global scale and a spatial resolution of about 1 km [16]. The average precipitation is calculated considering the time period from 1970 to 2000 and is published for each month individually (example is given in figure 9) [16]. Although precipitation data on a daily basis would better represent the precipitation trend, there is no reason to assume a better reference water map could be achieved. It can easily be integrated into a workflow because of the relatively small amount of memory which is needed to store the whole dataset, the usage of GeoTIFF as file format (which can be read by nearly every toolkit available) and no additional costs.

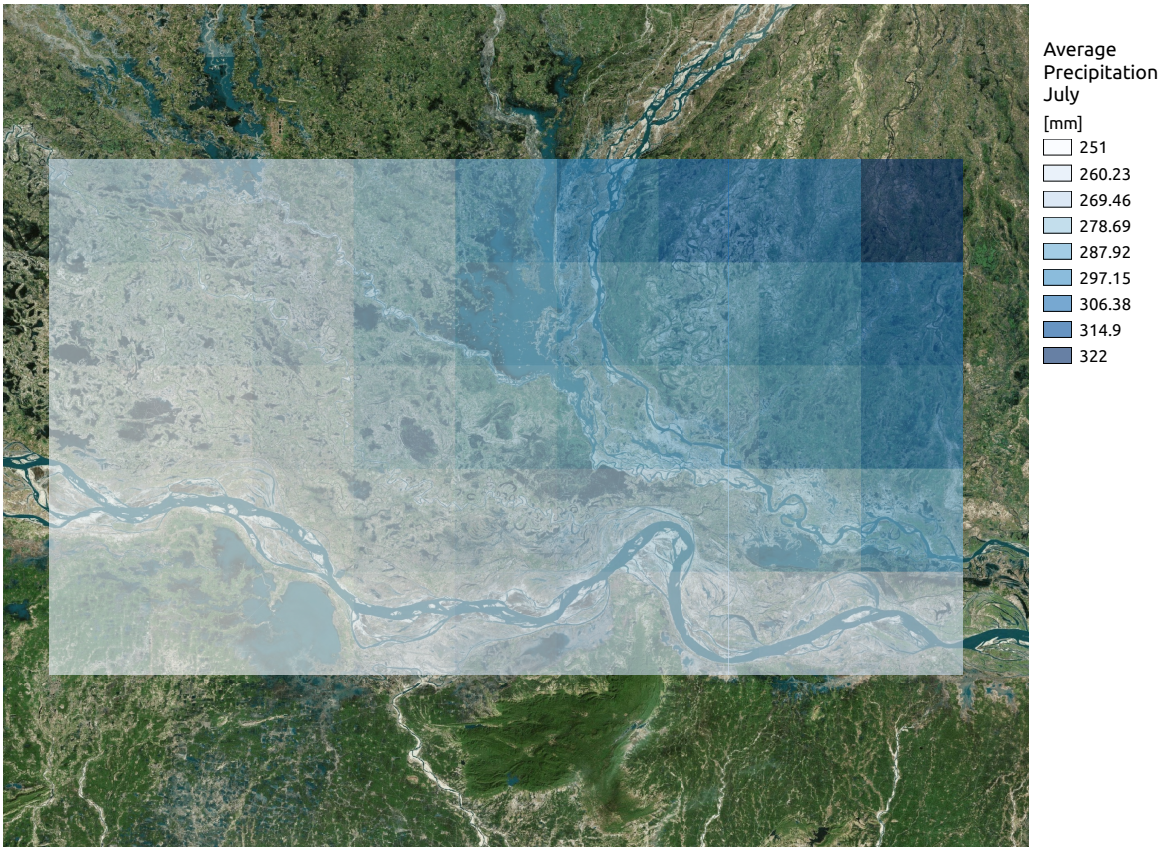


Figure 9: Average precipitation data of Bihar in July (data from WorldClim [16], Orthomosaic in background from BingMaps [1])

3.3 Tools

Python As described in [34], Python is an interpreted programming language which is well-designed and object oriented. It serves multiple roles since it can be used as procedural scripting or even functional programming language. One can not completely avoid the object oriented approach as it is implemented into the core modules of the language. Some reasons to use Python in geospatial analysis are listed here [34]:

- Computationally intensive tasks can be calculated efficiently due to the binding to C, C++ or other compiled-languages.
- It is a cross-platform language because of its connection to the C programming language.
- Geospatial analysis often focuses on real-world objects which can be easily modelled by the object-oriented approach of Python.
- Python is often used for application's APIs. For example Google Earth Engine platform provides a Python API.
- Many commonly used libraries (e.g.: GDAL, OGR, PROJ, GEOS) for geospatial analysis and in particular remote sensing can be accessed through Python.

Within this thesis, Python was used to script the algorithm which calculates reference information from an water map archive (for details see Section 4).

NumPy NumPy is written in particularly for Python and scientific computing but by the use of C programming language [34]. It is an extremely fast, multidimensional Python array processor which can interact with other important libraries [34]. Thanks to the interaction between Geospatial Data Abstraction Library (GDAL) and NumPy, large satellite images can be processed efficiently by the use of NumPy's capabilities.

The combination of different input datasets and retrieval of reference information for this thesis is done by the capabilities of NumPy. By the use of GDAL and NumPy the processing speed could be increased dramatically.

GDAL GDAL and its sister library OpenGIS Simple Features Reference Implementation (OGR) are responsible for an enormous amount of tasks within the geospatial industry. GDAL was created in the late 1990s by developer Frank Warmerdam and is mostly written in C or C++. Due to its C/C++ base it is focused on computational speed, which is important because of the large datasets geospatial industry has to deal with. GDAL's raster capability is part of nearly every geospatial toolkit (over 80 pieces of software) which includes very common ones like ArcGIS, Grass GIS, Orfeo ToolBox or QGIS [8, 34]. Some reasons for the library's success are [34]

- the X11/MIT open source license, which is both commercial and open source friendly.

- its multiple ways to access the capability. While there are simple ways to run GDAL through command line commands there exist a number of programming APIs for some of the most common languages (C, C++, Python and Java).

GDAL represents a substantial part of the thesis' algorithm. It is used for reading the input water maps, re-projecting all raster files to a common form, processing of the data in cooperation with NumPy and exporting the resulting images to the local file system.

QGIS QGIS (former Quantum GIS) is a professional, cross-platform and open-source GIS application [11]. It offers a well designed graphical user interface (GUI), a large number of geospatial tools and is able to handle many data formats [34]. QGIS has a Python integration, is based on GDAL/OGR and GEOS and can easily be automated [34].

Within this thesis it was used for basic GIS tasks, as flexible data viewer and for the production of resulting maps.

4 Methodology

As described in the objectives of this thesis (Section 1.2), a flexible work-flow for reference water map computation based on given water maps is designated as one result of this thesis. All tools, which were used for the underlying algorithm, were already described in Section 3.3. The basic work-flow is described in Section 4.1, while available parameters for possible variations are shown in Section 4.2.

4.1 Computation of a reference water map

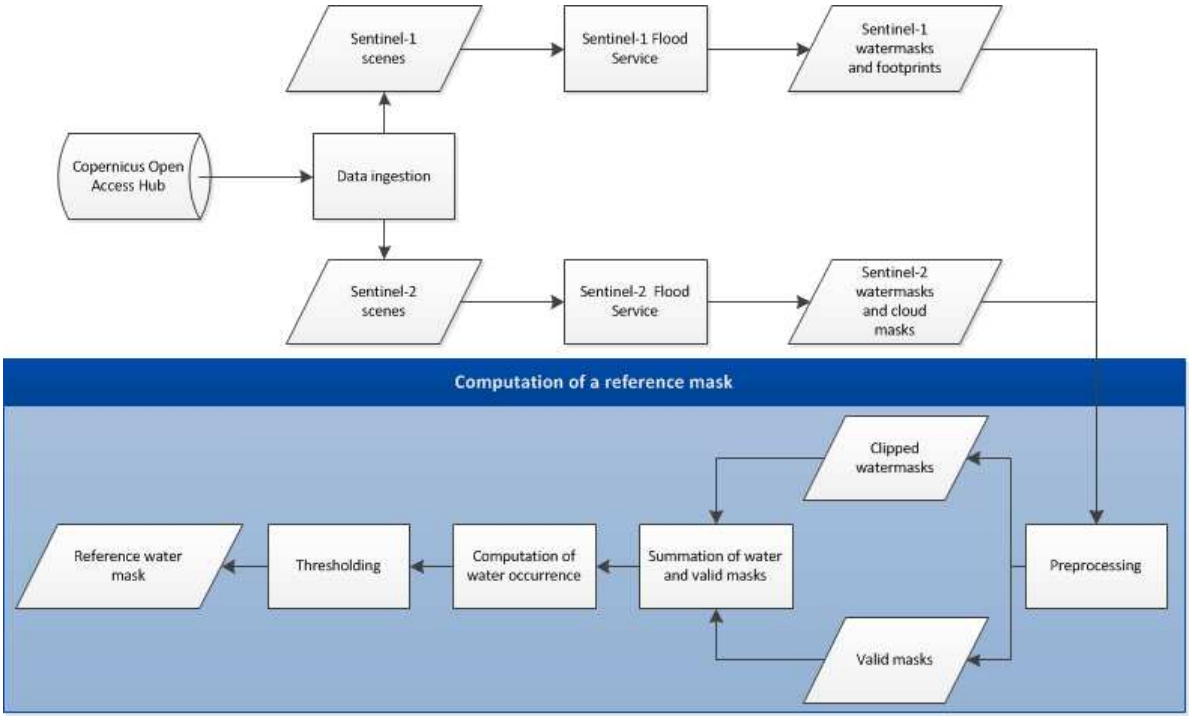


Figure 10: Overview of the methodology to process a reference water map based on Sentinel-1 and/or Sentinel-2 data

In [38] and [50], long time periods of optical data have been analysed to produce information products like water occurrence or permanent water areas. As described in [50], the water extent's temporal change and variation in space requires a distinction between permanent and seasonal water bodies. One common instrument for this distinction is called *frequency analysis* [50]. The frequency of a pixel to be classified as water in a given time period (=water frequency or occurrence) is analysed to decide whether the tested pixel belongs to the permanent or seasonal class (more details are given in Section 4.1.4) [38, 50]. Although the corresponding datasets (GSWE and G3WBM) were not primarily designed to serve as reference water map within the flood detection process, their *frequency analysis* approach is used as base for this work-flow. The original purpose of most surface water datsets like GSWE and G3WBM was to monitor changes of

the water areas in order to observe consequences of humane activity and natural changes [38].

An overview of the work-flow is given in Figure 10. The data ingestion and surface water detection (described in Section 4.1.1) is located in the upper half of Figure 10 and was performed by the already developed processing chains of the DLR [44, 47]. The second half of Figure 10 (blue background) describes the work-flow for calculating a reference water map.

4.1.1 Data ingestion and surface water detection

Based on a given area of interest and period of time, scenes observed by S-1 and/or S-2 are loaded from the Copernicus Open Access Hub [2]. Using the automatic processing chains 'S-1 Flood Service' (described in Section 3.2.1) and 'S-2 Flood Service' (described in Section 3.2.2) developed by the research group 'Natural hazards' in the department 'Geo-Risks and Civil security' of the DLR [4], the binary water maps are processed. Each water map represents the detected surface water of a certain observation which was done either by S-1 or S-2. Open surface water is marked as '1' while other pixels contain '0' as value. As by-product of the processing chains the footprint of the S-1 scene and cloud mask of the S-2 scene are computed. Both datasets are essential in the further processing.

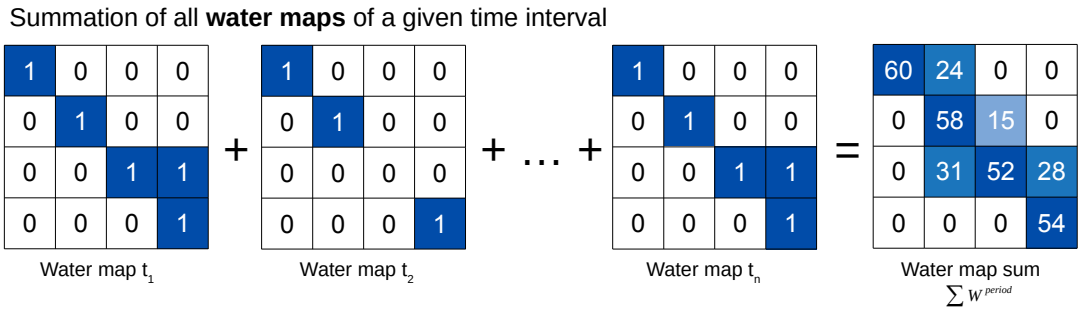
4.1.2 Preprocessing

To gain the ability to compare or combine S-1 and S-2 water maps matching projections, pixel sizes and spatial extents have to be ensured. In order to achieve this match, all water maps, footprints and cloud masks are transformed based on the specification of SWBD [13]. Therefore a transformation into WGS84 (EPSG-code: 4326) and a regular 1° grid is performed. As a consequence, every part of a water map can be assigned to a specific 1° cell and a common coordinate system is implemented. Clipping and re-projection is done by the use of GDAL's built-in *Warp* method which uses *nearest neighbour* re-sampling [7] by default. Because of the similar structure to SWBD the resulting data can be easily compared to available datasets and existing processing capabilities can be used. As a result of the preprocessing step clipped and transformed water maps and the corresponding valid masks (details are given in 4.1.3) are produced.

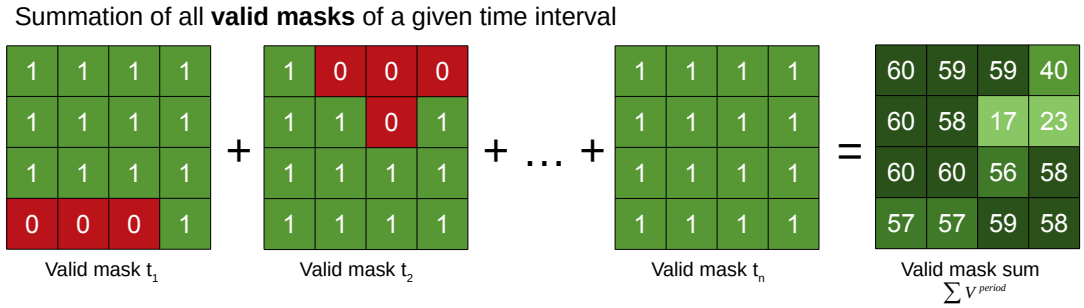
4.1.3 Summation of water and valid maps

Within this step, the frequency of water detections per pixel is counted within a given time period ($[t_1, t_2, ..., t_n]$). Therefore, the available water maps (in binary format) are summed up (Figure 11a) whereby the number of times a pixel is classified as water can be extracted ($=\sum W^{\text{period}}$).

[7] In the process of re-sampling an output image is defined by the desired scale, size and projection. The output image is transformed to match the the original image. Within the next step the value of the closest pixel (*nearest neighbour*) is assigned to each pixel of the output image. Due to the non-consideration of additional pixels the original intensity values will remain unchanged [20].



(a)



(b)

Figure 11: Exemplary presentation of the summation process step

To be able to quantify the percentage of water detections within a certain period of time, the maximum number of possible water detections per pixel (*valid observations*) has to be known. Defining this number, the exact area on ground which is captured by the satellite (*footprint*) has to defined. In case of optical observations like S-2, only pixels which were not covered by clouds or ice during the observations are valid for water detection. Due to the cloud penetrating abilities of radar observations, the entire area of the corresponding footprint is assumed to be valid. Further, the S-1 footprint data and S-2 cloud masks are transformed to binary valid masks. As a consequence, for each S-1 or S-2 observation a corresponding valid mask is available. Each pixel contains the information whether it would have been possible to detect water (1) or not (0). The general number of valid pixels within the given time period is computed by summing up all valid masks ($=\sum V^{\text{period}}$). An example of the described summation step is given in Figure 11b.

4.1.4 Computation of water occurrence

Water Occurrence (WO) is often used in literature to define the area which is permanently covered by water (see also: [38], [50]). As shown in Equation (4) [38], one can

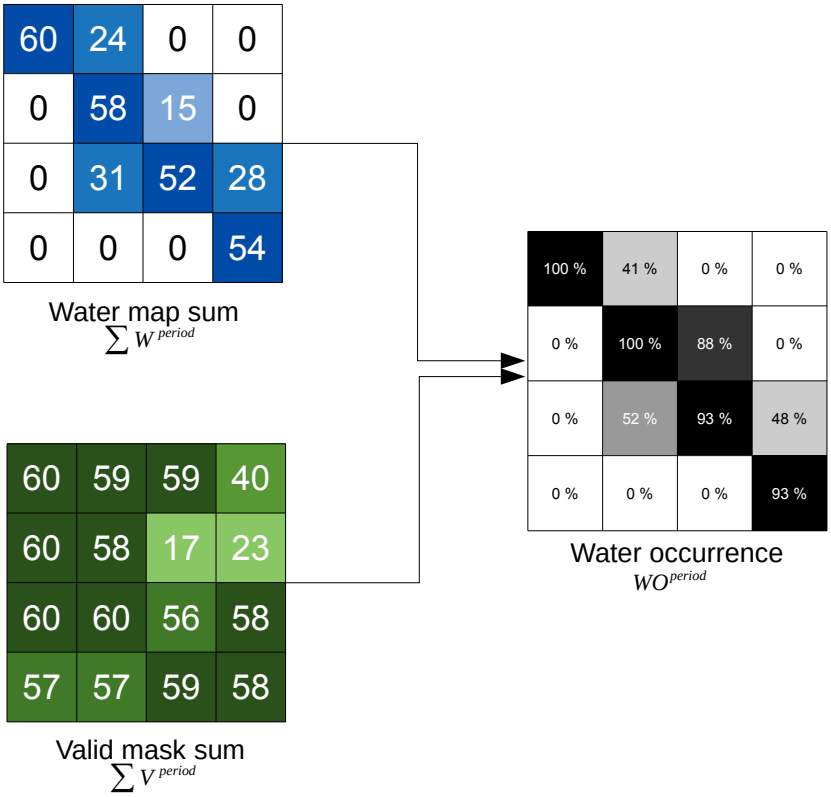


Figure 12: Exemplary presentation of the water occurrence calculation

calculate WO by dividing the number of water detections during a certain time period ($\sum W^{\text{period}}$) by the number of valid observations ($\sum V^{\text{period}}$) of the same period.

$$WO^{\text{period}} = \frac{\sum W^{\text{period}}}{\sum V^{\text{period}}} \tag{4}$$

Based on WO, which represents the percentage of water detections compared to valid observations, one can evaluate how frequently water was detected during a period of time (see Figure 12). In contrast to the number of water detections, WO can be interpreted independently of the number of input maps. This will be a condition of the next processing step.

4.1.5 Thresholding

By applying a threshold for WO (Equation (5) and Figure 13), those pixels which are assumed to be covered by water under normal conditions will be part of the resulting reference water map which is exported as binary image. Since there is a single threshold value selected for the whole map, it is called *image-wise threshold* in subsequent sections.

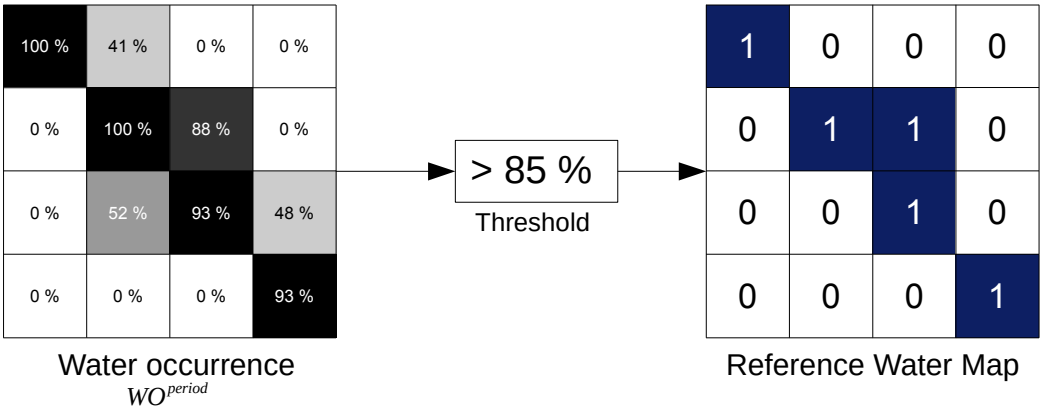


Figure 13: Exemplary presentation of the image-wise threshold application

Within this thesis, the threshold was determined based on visual inspection or by using an automatic approach which is presented in Section 4.2.4.

$$WO^{\text{period}} \geq \text{threshold} \tag{5}$$

In literature the image-wise WO threshold is used to distinguish between permanent and seasonal water surfaces. Therefore, literature provides various different WO threshold values e.g.: 100 % [38], 95 % [21], 90 % [19], 70 % [50] or even 50% [28, 40]. Most of the thresholds are defined by visual inspection of the resulting water bodies while others are determined by analysing the underlying frequency (e.g. by use of a 2D scatter-plot [21]). As mentioned (Section 1.1), there exist no ground-truth data of water bodies under normal conditions which means that there is no way yet to evaluate the quality of a defined WO threshold.

A truly permanent water area would mean that there was observed water coverage in every single observation of the considered time period (WO threshold of 100 %). The aim of reference mapping is to map the observed water bodies at normal conditions [29]. This implies that the impact of hydrological extreme events - like flood events or droughts - on the resulting reference water map has to be kept small. The presence of droughts causes that some pixels are temporally not classified as water which yields a decrease in WO. As a consequence, the affected pixels would not be part of the resulting reference map although they are inundated under normal conditions. Because of this, actual permanent water areas are seen to be inappropriate for reference water mapping. A similar effect as described for droughts could be caused by certain classification errors. On the other hand, a too low threshold could yield seasonal water surfaces to be falsely classified as part of the reference map [19]. As a compromise, areas which show water occurrence of more than 90 % are considered as part of the reference water map, like it was stated in [19]. This threshold suggestion was used as a reference value and was

altered according to local conditions. In [29] it is mentioned that there exist numerous definitions what are considered as *good reference*. Because of this, threshold values can not be perfectly optimised.

4.2 Parameter variation

Since the variability of water areas vary widely on global scale [38] and to achieve the best possible result, the described method (Section 4.1) has to be adapted locally. Subsequently, some parameters of adaptation are presented and their effects on the final results are analysed later in Section 5.1.

4.2.1 Considered period of time

In order to get reasonable reference information, a specific period of time is selected and only water maps which were observed within this period are considered. In the scope of this thesis, distinct years are considered as maximum time periods because of the immense processing time of a multi-year calculation. In detail, the year 2017 and 2018 (for Bihar) were used in the calculations and the difference of both years is investigated in Section 5.1.2. Within one selected year it has to be decided whether all available maps will be considered or a well-defined subset. Depending on the hydrological region, the considered period of time has an important impact on the final reference map. Especially within subtropical regions the temporal distribution of valid optical observation due to rainy season or dry periods can be altered by this parameter. The lower the number of water maps, which are considered during the calculation of a reference water map, the larger the impact of single water maps will become. Since some water maps could be affected by sensor failures or processing errors, this impact of the single water map determines the risk of errors within the final result. In general it can be controlled whether extreme events are included in the calculation, the temporal distribution of optical observations and how error-prone the result will be.

To limit the considered period of time to a time slot, which represents water bodies under normal conditions, two values are taken into account: Water cover and average precipitation. Water cover (W_{cover}) is the percentage of water pixels (n_{water}) compared to the overall number of pixels within a certain grid cell (n_{overall}).

$$W_{\text{cover}} = \frac{n_{\text{water}}}{n_{\text{overall}}} \tag{6}$$

Precipitation data is obtained by WorldClim as described in Section 3.2.3. As shown in [40], average precipitation and water area might be correlated. Due to the fact that precipitation is not the only factor influencing the water area, it depends on the regional hydrologic system if it can be considered a valid value for defining the observed period of time.

Within this thesis the following options are compared, which are defining the considered period of time:

Entire period The whole time period will be included and all available water maps of the selected year are used for calculation. Every hydrological extreme event will have an impact on the final result.

Months of median water cover (MMWC) For each monthly water map water cover (see Equation (6)) is calculated and a yearly median water cover is processed. The considered period of time consists of the months whose water cover is closest to median water cover, in the following referred to as months of median water cover (MMWC). On the one hand the processing time can not be decreased since every water map has to be analysed. On the other hand the impact of flood and drought events on the resulting reference water map can be limited. If an up-to-date reference water map is maintained continuously the actual processing time might be less important and the median water cover approach offers a reasonable alternative.

Months of median average precipitation (MMAP) By the use of regional average precipitation data (for details see Section 3.2.3), which are provided on a monthly basis, a yearly median precipitation value is calculated. An option for the considered period of time is to select the months whose average precipitation is closest to median average precipitation, in the following referred to as months of median average precipitation (MMAP). If average precipitation and water extent are correlated the impact of extreme events can be limited as well. An advantage of the median precipitation approach is the use of easy available precipitation data. Because of this, processing time could be reduced drastically as a smaller number of water maps has to be considered. The drawback is that an analysis has to be done which evaluates the correlation between precipitation and water extent.

4.2.2 Sensors

The work-flow of Section 4.1 can be executed using water maps from a single sensor or from multiple sensors. As shown in [33] and [39], the shared usage of multiple datasets and especially radar and optical sources allows to benefit from both sensor's characteristics. According to [33], the joint and systematic use of S-1 and S-2 will be improving the mapping of open water bodies in future. The high revisit rates, good spatial resolution and synergies between SAR and optical data are mentioned as further advantages. As the developed work-flow is intended for flood detection purposes, only a narrow time slot is available to gather satellite data. In practice, only one sensor will be available during this time slot and due to the frequent need of cloud independent observations, often radar data will have to be used. As a consequence, the usage of reference water maps from a combination of multiple sensors would have to be compared to single sensor products in flood detection. Because of this, the resulting flooded areas would still be affected by sensor specific limitations which is the reason why this thesis is focused on the differences between multiple sensors to improve flood detection.

Since all water maps and valid masks are delivered in binary format, they can be treated identically and independently from the used sensor in terms of the algorithm.

As a consequence, the work-flow can be executed by the use of S-1 or S-2 water maps without sensor specific adaption. If water maps of another sensor would have to be supplemented (for example in case of missing observations during the rainy season), they can be easily inserted into the input data. In theory every sensor, which could be used for water detection and delivers binary water maps, can be included into the calculation.

4.2.3 Monthly normalisation

Most datasets, which represent permanent water areas, are produced by direct summation of the single water maps of a certain period (for example see: [21], [50]). In [38], an approach which makes use of monthly water maps is presented instead. As intermediate step, reference water maps are calculated in the same manner as presented in Section 4.1 for each month within the selected period of time. Further, WO is calculated by the sum of these monthly reference water maps instead of single water maps. On the one hand the same impact of every month of the year can be ensured by introducing this intermediate step. However, as mentioned in Section 4.2.1, the error rate increases due to a smaller number of scenes defining each monthly water map. Due to cloud cover, volcanic eruptions or other extreme events especially optical observations could be problematic. Subsequently, the assumption of the water occurrence could be distorted. If the number of scenes within some months is noticeable low, they could be replaced by water maps of another sensor as it is suggested in Section 4.2.2.

4.2.4 Automatic or pre-defined threshold

As mentioned in Section 4.1.5, the WO threshold is applied to decide whether a pixel is part of the reference water maps or not. Additionally, the image-wise threshold was presented and examples from literature were listed. The image-wise threshold has to be defined manually prior to the actual processing based on visual inspection, expert knowledge or a common value in literature. In this case each pixel of a scene is compared to the same threshold value (as shown in Equation (5)).

Alternatively, an automatic or pixel-wise threshold is presented in this thesis. The pixel-wise threshold value is calculated using a measure for the dynamic behaviour of the corresponding pixel. The idea is based on the fact that reference water maps should represent water areas under normal conditions. Achieving this condition is especially challenging in hydrolgical dynamic regions, where normal conditions might be present for a shorter period of time, compared to hydrological more static regions, due to a larger number of hydrological extreme events and more variations of water extent. Additionally, water bodies could be altered geographically during one year which further decreases the WO. Based on this observations, the pixel-wise threshold will be lower for pixels which show a more dynamic behaviour over time and it will be larger for pixels which show a more static behaviour.

The calculation of the automatic or pixel-wise threshold is given by

$$\text{threshold}_{\text{pixel}} = (1 - \nu_{\text{dynamic}}) - 0.1, \tag{7}$$

$$\nu_{\text{dynamic}} = \frac{n_{\text{changes}}}{n_{\text{valid}}}, \tag{8}$$

where the term $\nu_{dynamic}$ is explained by Equation (8). A single change appears if the pixel value of the previous time step is different compared to the current one. The number of changes n_{changes} is divided by the total number of valid observations n_{valid} within the considered period of time and for each pixel. As a consequence, $\nu_{dynamic}$ will be a number between 0 and 1, where pixels, which show a more dynamic behaviour, are represented by a higher number. To achieve the circumstance, that more dynamic pixels will be compared to a lower threshold, $\nu_{dynamic}$ is subtracted from 1 in Equation (7). As mentioned in Section 4.1.5, a threshold of 90 % is suggested in this thesis and a threshold of 100 % should be avoided. Based on this conditions, the pixel-wise threshold is reduced by 10 % (or 0.1) in Equation (7), to limit the maximal possible threshold by 90 %. The pixel-wise threshold is applied by comparing each pixel to its individual threshold $\text{threshold}_{\text{pixel}}$ (see Equation (5)).

Beside the consideration of the dynamic behaviour of a pixel, an additional advantage of the pixel-wise approach is the plausible and objective threshold definition which avoids manual regional adaption. If comparability between reference maps of different areas of interest is on demand the image-wise approach should be selected. The same is also true for the need of manual choice of threshold to consider expert knowledge or specific local circumstances.

5 Results

The work-flow described in Section 4.1 can be influenced by the parameters presented in Section 4.2. As a final product of the work-flow a reference water map is produced. In comparison to a single water map, which represents the water bodies of a satellite observation at a specific date, the reference water map represents water bodies under normal conditions.

In Section 5.1, the resulting reference water maps are compared to each other to monitor the effects of certain parameters. To display different properties of two reference water maps, a difference map is calculated. Using this approach, pixels which are only classified as water by one of the maps and pixels which are classified as water by both maps are coloured differently.

In Section 5.2, the reference water maps generated in the scope of this thesis are compared to already available global surface water products. In detail SWBD and multiple versions of GSWE are considered for comparison purposes.

As a completion of this thesis, a flood detection process is simulated in Section 5.3. It is shown how the information from Section 5.1 can be used to better understand the characteristics of reference water maps and to improve the flood detection process.

5.1 Effects of parameters on the final reference water map

The configuration of parameters, which are presented in Section 4.2 (considered period of time, sensors, monthly normalisation and threshold), can have an influence on the resulting reference water map. Within this Section, the effects of theses parameters should be investigated.

5.1.1 Comparison of Sentinel-1 and -2 reference water maps

In this Section, the impact of the selection of a sensor type is investigated. While S-1 delivers radar based data, S-2 acquires optical images. For the reason of distinct processing chains and sensor specific characteristics the resulting reference water maps do not appear identically. To be able to draw reasonable conclusions, e.g. for flood detection, one has to know which sensor was used in processing and which limitations should be considered. The reference information's sensor-specific characteristics can be observed by comparing the corresponding results and detect diverging properties. In practice, the availability of a certain sensor is the crucial factor for sensor selection. As mentioned in Section 4.2.2, this is particularly true in case of time-critical applications like flood detection.

In case of reference water mapping the knowledge of the sensor-specific characteristics is therefore a reasonable method to discover potential error sources within the resulting map. In further explanations the reference water maps processed from S-1 water maps only are called S-1 reference maps and water maps processed from S-2 water maps only are called S-2 reference maps.

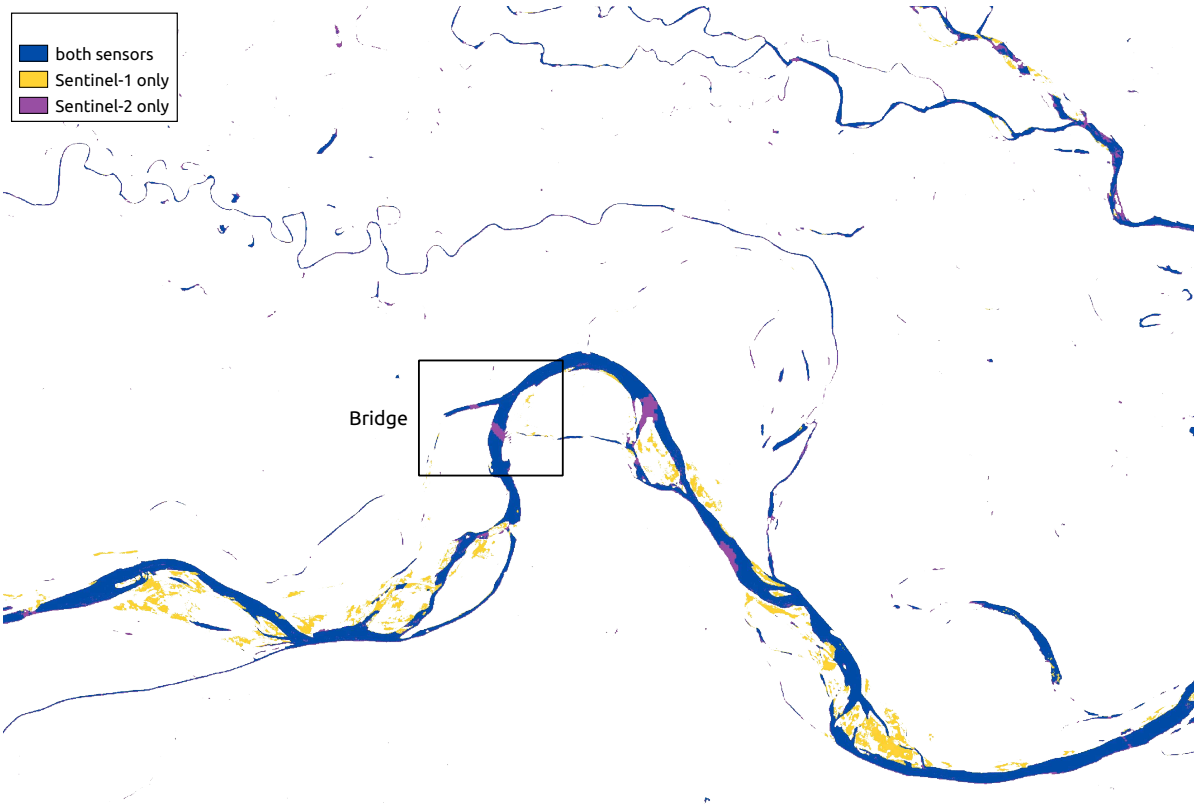


Figure 14: Comparison of S-1 and S-2 reference maps in Bihar

Figure 14 shows the differences between a S-1 and S-2 reference water maps. The reference water maps were calculated using the entire period of water maps of the year 2018, without a monthly normalisation step and by applying a image-wise threshold of 85 %. One will notice some areas which are only observed by S-1 (coloured in yellow). On the other hand, there are some areas which are only included within the S-2 reference water map (coloured in violet). Areas, where both maps agree by classifying the corresponding pixels as part of the reference water map, are coloured in blue.

As mentioned in Section 2.1.2, the backscatter behaviour of surface water and other land cover categories (like wet sand or ice) could be the same and subsequently, the categories could be confused. Within the presented scene the river is framed by sand which leads to an overestimation of the S-1 water detection. In Figure 15, a temporally close S-2 image is compared to the detected water areas of S-1. While the grey area represents the water area observed by S-2, the corresponding S-1 water detection is coloured in white. One can observe that the surrounding sandbank is irregularly classified as water. The multi-spectral data observed by S-2 is not affected by this confusion, which is the reason for the detected differences. This kind of overestimation of radar based water detection is also described in [19]. In [33], land cover data are used to avoid this type of confusion.

The (violet coloured) areas in Figure 14, which are only observed by S-2, can be

Figure 15: Example for S-1's overestimation of the water extent

explained by multiple reasons. The linear-shaped area, which is highlighted in Figure 14, is caused by a local bridge. In detail it refers to the *Munger Ganga bridge*, which crosses the Ganges river and has a length of 3.7 kilometres. The double-bounce scattering, which is caused by multiple reflections of the radar radiation between water and bridge, increases the backscatter value locally. Although the bridge does not cover a whole pixel in the corresponding water maps, the backscatter increase still affects the pixel value. Since only low backscatter values are determined as water, the S-1 processing chain excluded the affected pixels from the result. The S-2 based water detection on the other hand seems not to be affected in the same manner.

One step of the S-1 processing chain is the application of an exclusion mask which should reduce potential overestimation (for more details see section 3.2.1). Figure 16 shows an overlay of Figure 14 and the mentioned exclusion mask. It can be noticed that some of the areas, which are only observed by S-2 can be explained by obvious uncertainties within the used exclusion mask (coloured in dark violet). In case of further improvement uncertainties of the exclusion mask could be identified by the comparison of S-1 and S-2 reference water maps.

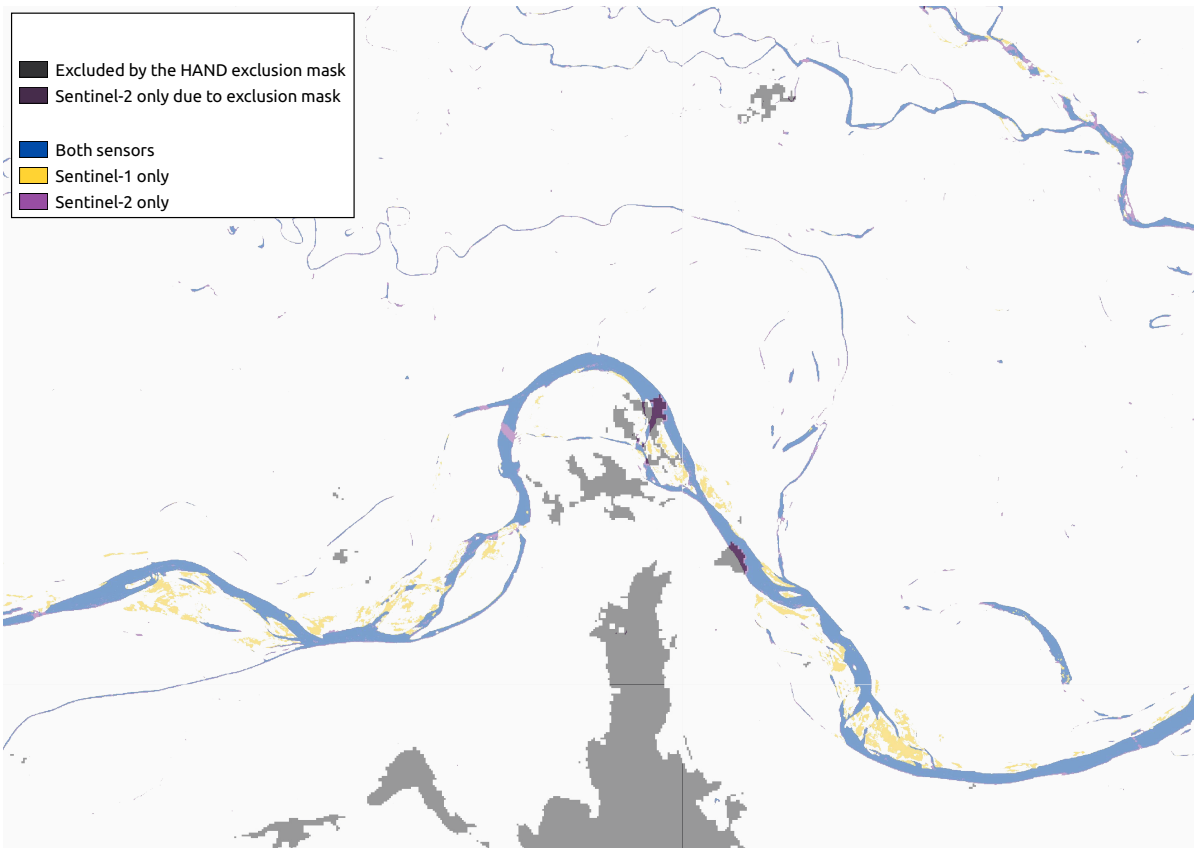


Figure 16: Difference of S-1 and S-2 reference maps compared to exclusion mask

The observed scene shows many small inflows of the Ganges river especially in the northern parts of the scene. One can notice that the detected water areas of these river meanders differ between S-1 and S-2. The S-2 reference water map shows a more complete representation of these narrower inflows. Both processing chains are meant to observe large water bodies and they are not particularly optimised to represent narrow rivers. The resulting water maps of both processing chains are delivered in the same spatial resolution of 30 m. Since water waves increase surface roughness, the backscatter value increases as well [43]. Water waves can be induced by winds [43] or river turbulence can be responsible for an increase in surface roughness. As described in Section 3.2.1, within the S-1 processing chain a fuzzy-logic-based classification refinement is performed. One criterion to define the degree of water class membership is the size of the potential water area. As a consequence, smaller water areas are less likely to be part of the water class and could be classified as non-water area. Due to the side-looking aperture of SAR, objects which are located on the river bank (e.g. vegetation) could influence or cover the actual river [43]. This is especially problematic in case of narrow rivers, whose width is represented by a low number of pixels. These characteristics are assumed to be the reasons for the gaps which appear within the S-1 reference water map regarding narrow river meanders.

Figure 17 shows a comparison of S-1 and S-2 reference water maps in Zaragoza. As

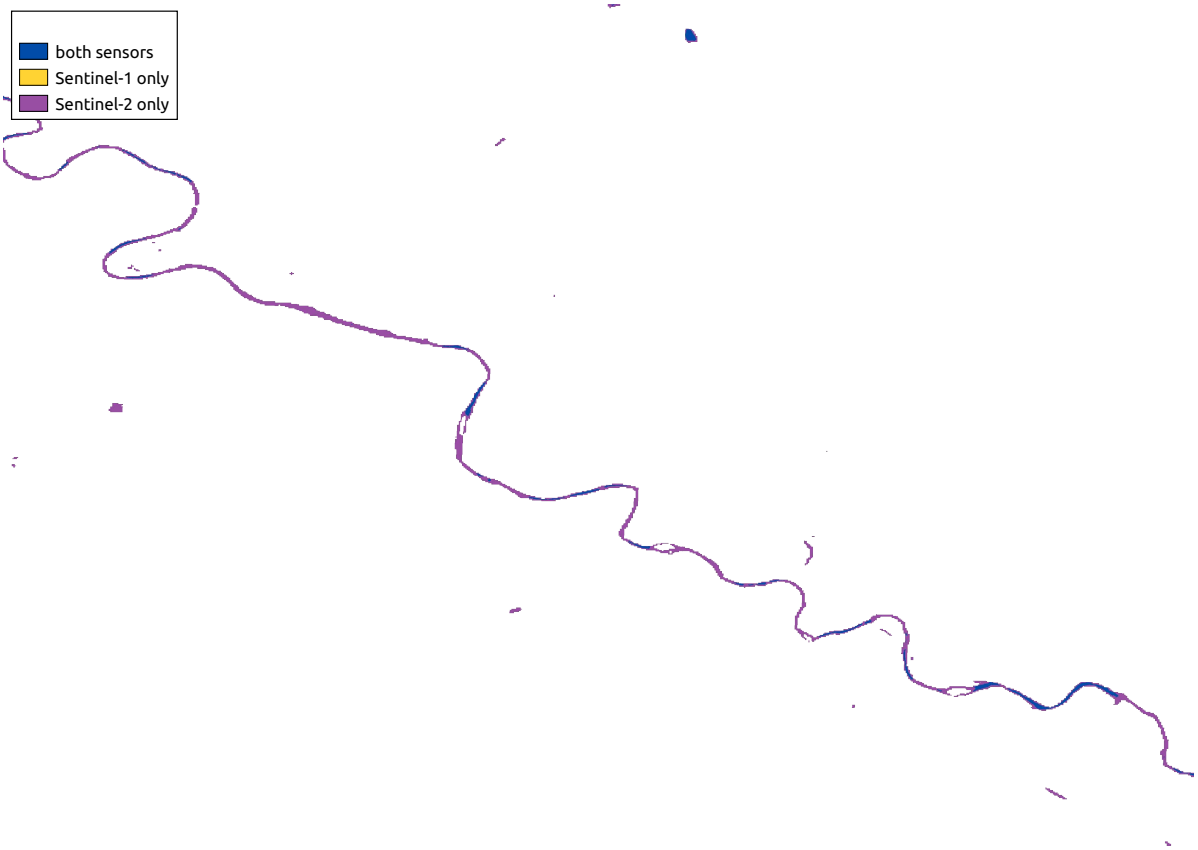


Figure 17: Comparison of S-1 and S-2 reference maps in Zaragoza

mentioned within the description of the area of interest (Section 3.1), the selected area surrounding Zaragoza differs in hydrological and climatic behaviour compared to India's Bihar. Additionally, the river Elbro is narrower compared to Ganges river. One can notice similar results when comparing Bihar's narrower river meanders to Zaragoza's river situation. Analogous to Bihar, S-2's reference water map provides a more complete representation of the local river system. It is noticeable that the gaps within the S-1 reference water map are mainly located near the river bends or bottlenecks of the river. In the Figures 18 and 19, some exemplary regions are highlighted to investigate the gaps within the S-1 reference water map. While Figure 18 shows a S-2 true-colour image of a selected subarea of the are of interest in Zaragoza, the corresponding WO of S-1 in 2018 can be seen in Figure 19. It can be noticed that all regions of lower WO (light blue coloured areas in Figure 19) are located at river's most narrow parts (compare to Figure 18). This effect is assumed to be caused by the increased occurrence of waves in these areas which consequently increases the local backscatter value. Additionally, the river is occasionally surrounded by vegetation which is especially influential in the case of a side-looking aperture and a smaller river width. Another missing part in the S-1 reference map is the urban area of Zaragoza which can be result from the double bounce effect. The double bounce effect often occurs because of the city's buildings and majorly increases the backscatter value locally. Additionally, it can be noticed that the

overestimation of S-1 at the riverbanks, which is observable in Bihar, does not appear in Zaragoza. It can be assumed that the diverging soil conditions of both areas are the reason for this effect.

Figure 18: Analysis of S-1's incompleteness in Zaragoza (S-2 RGB image)

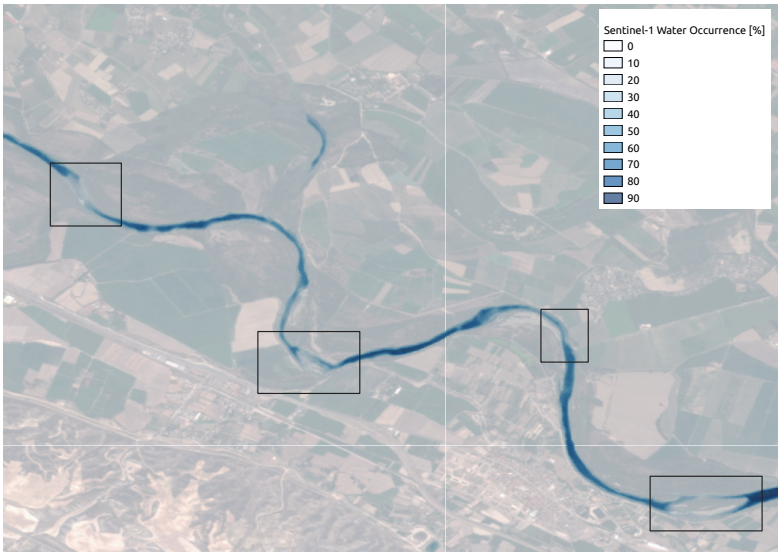
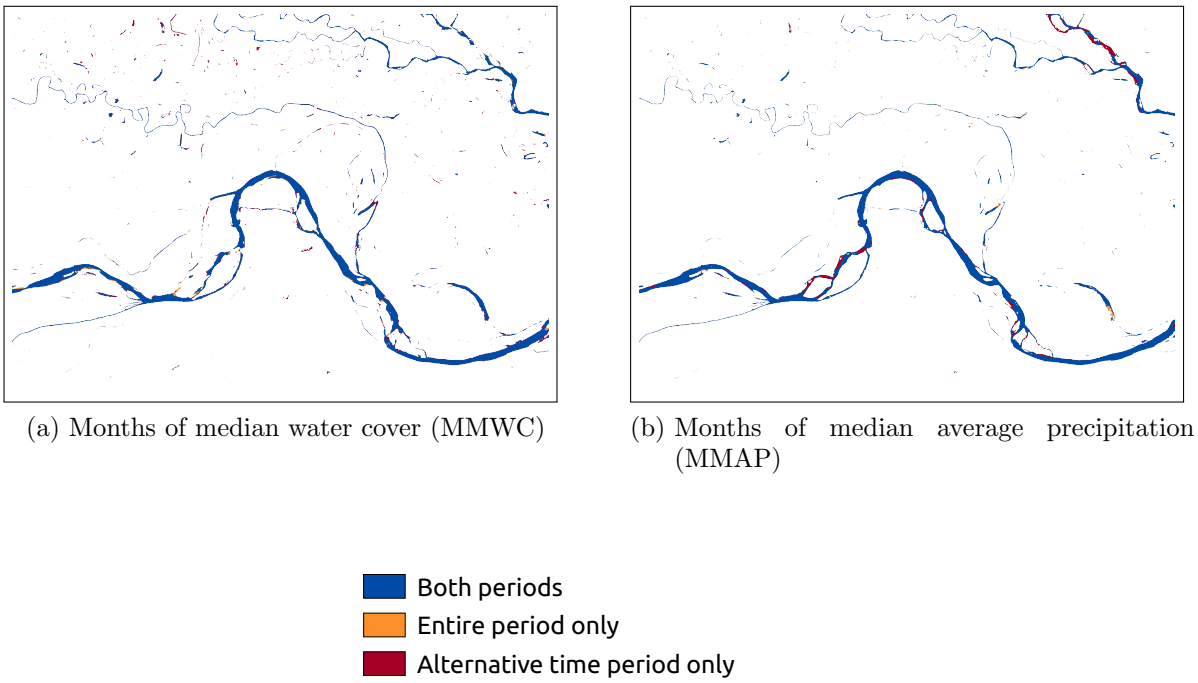


Figure 19: Analysis of S-1's incompleteness in Zaragoza (S-1 WO)

5.1.2 Impact of the considered time period

In Figure 20, S-2 the reference water map, which considers the entire period of the year 2018, without monthly normalisation and by applying an image-wise threshold of 85 %, is compared to alternative options for defining a considered time period (median precipitation and median water cover). The Figures 20a and 21b show difference maps between the reference water maps of the entire period and from one of the alternative option. Therefore areas, which are only part of the entire period's map are coloured in orange, while area, which are only part of the alternative option's map are coloured in red.



(a) Months of median water cover (MMWC)

(b) Months of median average precipitation (MMAP)

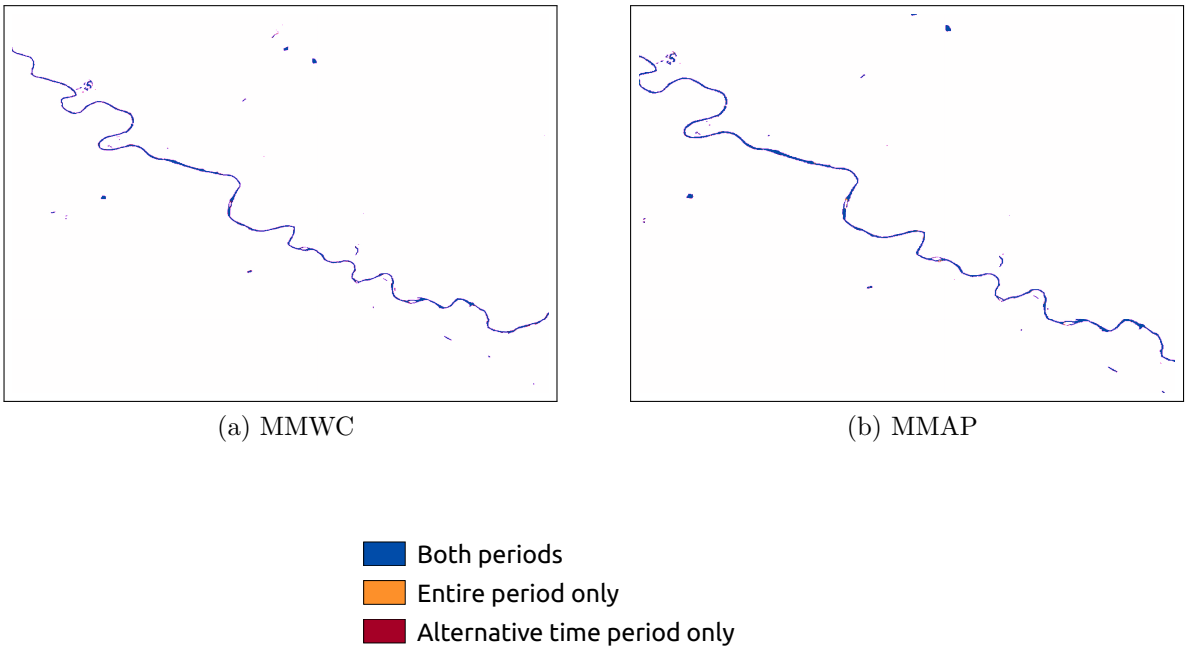
Figure 20: Comparison of multiple definitions of the considered time period in Bihar

One can notice that Figure 20a displays the difference map of the Months of median water cover (MMWC) and shows the most similar result compared to the reference map of the entire period. Figure 21b shows the difference map of the Months of median average precipitation (MMAP) and a more complete representation of the observed water bodies compared to the map of the entire period is noticeable. Although both approaches are using the same number of months their results differ obviously. The MMWC consist of January, February, November and December, while the MMAP contain January, March, April and May. Since there is no ground-truth data available, the time periods have to be compared by their definition process. Due to the fact that different months were selected by the use of median precipitation and median water cover, it can be assumed that precipitation and water cover are not correlated in this situation. The

reason might be that the amount of water in the area of Bihar is normally controlled by the precipitation of the Himalaya region in the north of Bihar (details of the hydrological situation are given in section 3.1.1). In general, it can be noticed that reference water maps from shorter time periods than one year could still show similar results. Because of the shorter period less variability between the participating water maps appears and a more complete image can be expected.

Although the MMAP yield the most complete representation of the local water bodies it is not recommended to select this approach in general. Due to the more complete image of shorter time periods it is recommended to use the median water cover approach for the region of Bihar.

In Figure 21, similar comparisons are done for the region surrounding Zaragoza. It can be seen that there are less differences between a shorter period of time and a whole year compared to Bihar. MMWC, MMAP and the entire period show similar results. This finding is caused by the more static hydrological behaviour within the region surrounding Zaragoza. As a consequence, even a shorter time period can be selected and processing time could be decreased without reducing the result's quality significantly.



(a) MMWC (b) MMAP

Figure 21: Comparison of multiple definitions of the considered time period in Zaragoza

Besides the impact of different options to select a considered time period wihtin a certain year, the similarity of reference water maps of two different years will be investigated as well. Highly dynamical regions like Bihar require up-to-date input data to represent the water bodies under normal conditions accurately. Water maps of multiple years differ in water extent and location of the river branches. An example of the water

cover over time is given in Figure 22. A similar trend can be observed for both years which is driven by the climatic seasons in the subtropical region of Bihar. One can still notice an obvious distinction between those two years which are noticeable in Figure 23 as well.

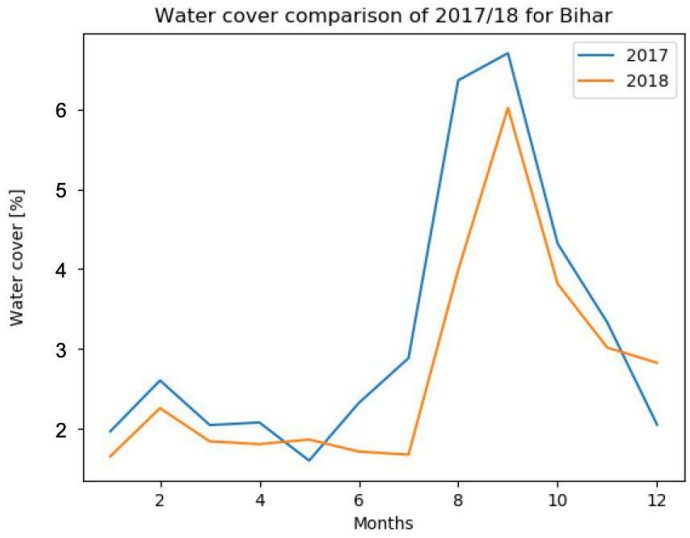


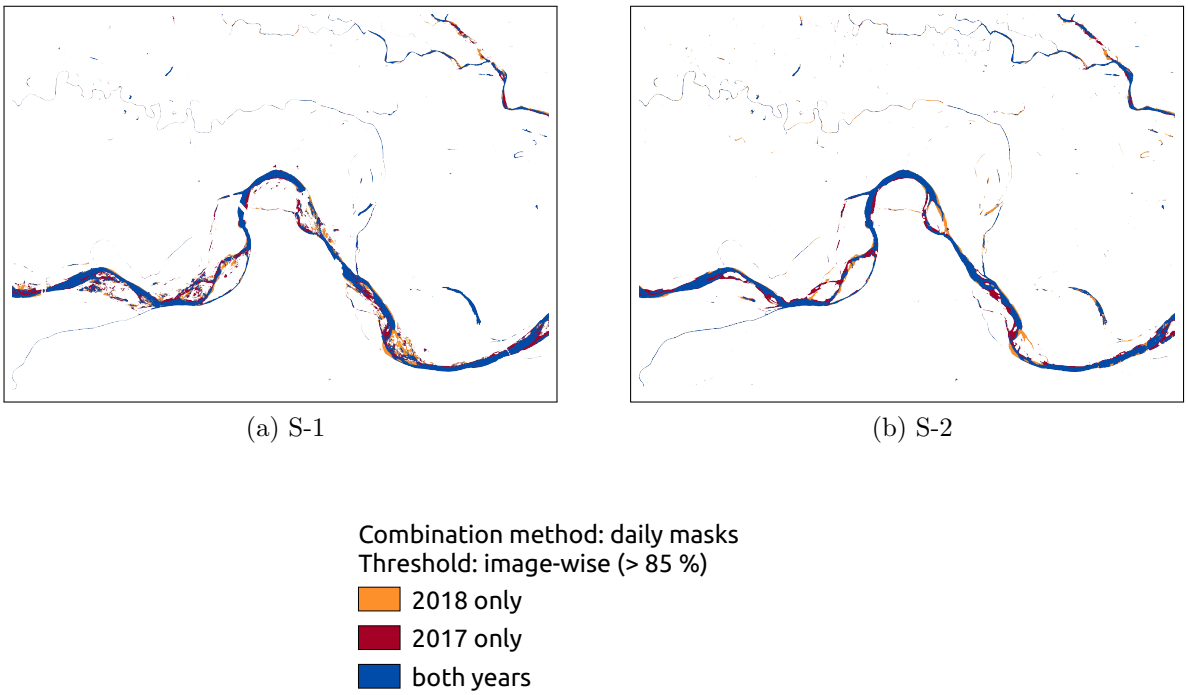
Figure 22: Comparison of water cover in 2017 and 2018 in Bihar

There, the reference water maps of 2017 and 2018 are compared for the region of Bihar. Therefors maps from S-1 (Figure 23a) and S-2 (Figure 23b) are considered. It can be seen that the differences of the S-1 reference water maps are mainly located in those regions where overestimation is assumed. Those gaps, which are explained by the exclusion mask used within the S-1 processing chain (for details see Section 5.1.1), exist in both years. This fact supports the assumption that the exclusion mask is the reason for these gaps.

Most differences between S-2's reference water maps (Figure 23b) can be located in narrower river parts. Since narrower river meanders are more often affected by spatial changes, this result seems to be realistic. If reference information of a inadequate year would be used, those narrow river parts would have been represented inaccurately within flood detection.

The effect of an additional S-2 satellite, which was launched on 7th March 2017 [12], cannot be seen within the corresponding difference map (Figure 23b). When observing the number of considered water maps of both sensors in Table 4 the launch is clearly noticeable. Due to this larger number of S-2 observations, the number of S-2 water maps, which are filtered by cloud coverage in the processing chain, increased as well.

Due to obvious differences between 2017 and 2018 in Bihar, it can be confirmed that up-to-date reference water maps are crucial for the flood detection's quality. In Section 5.2, where the thesis' reference water maps are compared to available surface water products which are partly not up-to-date either.



(a) S-1 (b) S-2

Figure 23: Reference water map comparison between year 2018 and 2017 in Bihar

	S-1 water maps	**S-2 water maps**
2017	67	58
2018	56	134

Table 4: Number of considered water maps in 2017 and 2018

5.1.3 Effect of the monthly normalisation step

In literature, most surface water bodies are directly built by the calculation of water occurrence from single water maps. Global Surface Water Explorer (GSWE) [38] is produced by an additional monthly normalisation step. In Figure 24, a comparison of two reference water maps shown whereby one was produced by combining each map directly and one was produced by processing monthly water maps first. Both reference water maps were processed considering S-2 water maps of the entire period of 2018 and by applying an image-wise threshold of 85 %. One can notice a more continuous representation of water bodies without the use of monthly normalisation (blue and orange areas in Figure 24). In the case where reference information is created from monthly water maps, a threshold is applied twice. First, when the monthly water maps are created and the second, when the monthly maps are combined to build the resulting reference water map. The doubled application of the threshold is assumed to be the reason of the incompleteness. As described in the following section, the pixel-wise threshold is able to improve the result created by monthly water maps. One disadvantage of the monthly

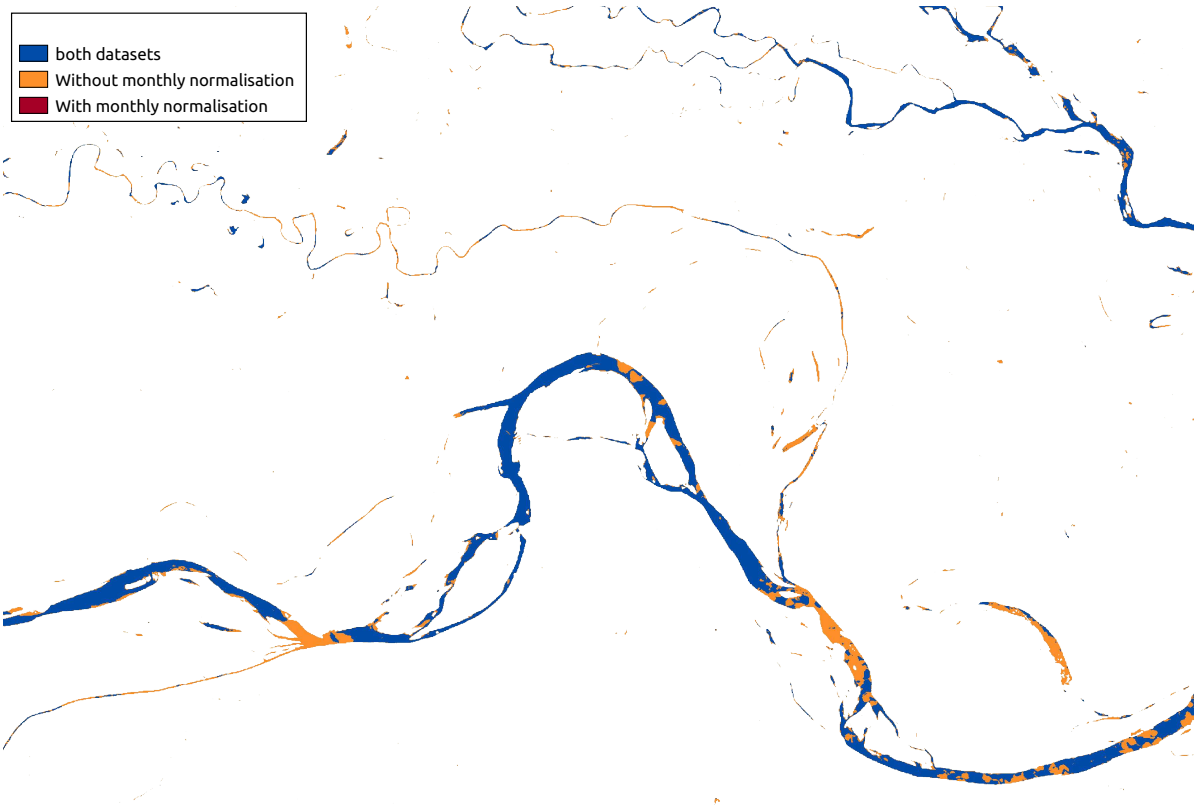


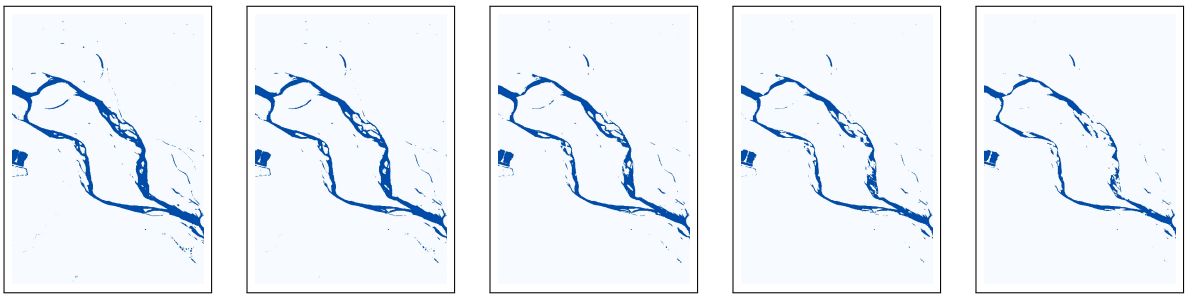
Figure 24: Difference of reference water maps which are produced by intermediary calculating monthly maps or directly combining daily maps

approach is an increased susceptibility to errors caused by the short time period of a single month. Due to weather extremes, sensor failures or processing errors individual months could consist of a low number of available input maps. As a consequence, the impact of a single water map increases and the monthly water map could be distorted more easily. When considering a threshold of 90 % (equals 10.8 months out of 12), a water pixel has to be part of 11 out of 12 monthly masks to be part of the final reference water map. By implication, already two distorted months could majorly affect the final result.

Because of the calculation of monthly water maps, it can be ensured that the water maps have an equal temporal distribution over the whole year. Due to fewer valid S-2 observations during the rainy season, a lower number of water maps is considered during this time. Although this might not have a visible impact on the final reference water map, it has to be considered when the result's quality is evaluated.

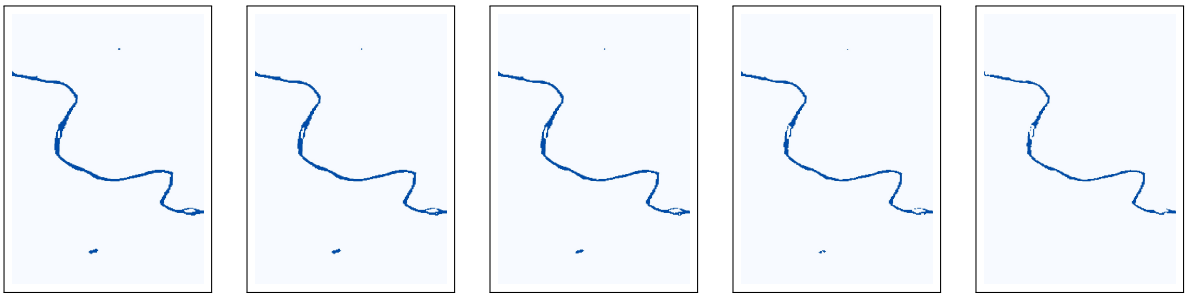
5.1.4 Impact of the selected threshold value

The threshold value defines from which WO level a pixel is seen as part of the reference water map. As already mentioned, there are no ground-truth data available for water



(a) WO > 50 % (b) WO > 60 % (c) WO > 70 % (d) WO > 80 % (e) WO > 90 %

Figure 25: Comparison of multiple thresholds values from 50 to 90 % in Bihar



(a) WO > 50 % (b) WO > 60 % (c) WO > 70 % (d) WO > 80 % (e) WO > 90 %

Figure 26: Comparison of multiple thresholds values from 50 to 90 % in Zaragoza

bodies under normal conditions. Due to this lack of information, it is hardly possible to optimise the threshold value. Alternatively, the threshold value is set by visual interpretation. The effect of different threshold values can be seen in the Figures 25 and 26. For these comparisons a S-2 reference water map was calculated by considering the entire period of the year 2018, without a monthly normalisation step and by varying the threshold value between 50 % and 90 %. It is obvious that the water extent decreases as the threshold value increases. While an obvious change between different threshold values can be seen in Bihar (Figure 25), the results in Zaragoza (Figure 26) appear more similar. These diverging effects are mainly caused by the different hydrological condition of both area of interests.

In [19], a compromise threshold of 90 % is suggested to compare different surface water products. However, as already stated, this threshold value was considered as reference value and was adapted locally to meet regional conditions. In case of Bihar the threshold reference value was reduced to 85 % in order to better work with the high dynamic situation. To improve the incomplete S-1 reference water map in Zaragoza, the local threshold value was reduced even further and was set to 70 %, although it is a hydrological static region. Actually, it could be assumed that a significantly larger threshold could be set in hydrologically similar regions, if not for the improvement of S-1 data.

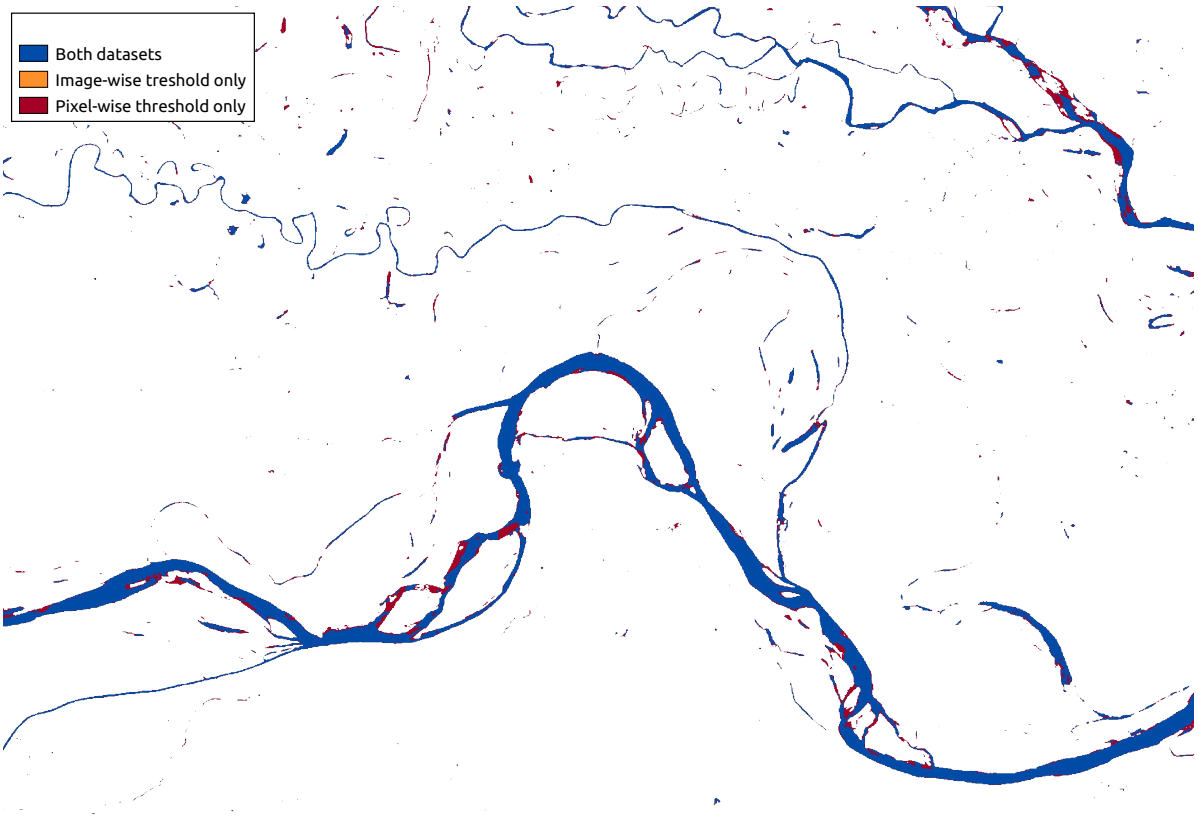


Figure 27: Difference of reference water maps which are produced by applying an image-wise or an automatic pixel-wise threshold

In Section 4.2.4, two types of thresholds are presented and their advantages and disadvantages are described. In Figure 27, the comparison between both threshold types is shown. It is obvious that the pixel-wise threshold defines a larger water extent compared to the image-wise threshold (> 85 %). On the other hand, a more continuous image can be produced and no manual definition of an image-wise threshold has to be done. It has to be mentioned that the average pixel-wise threshold of the whole scene is lower than the selected image-wise threshold for both areas of interest. This finding might be the main reason for the larger water extent in case of the pixel-wise threshold. If a pixel varies strongly during the year, the threshold will be lower compared to a more static pixel. This effect increases the number of months a pixel can be classified as non-water area and nevertheless be part of the resulting reference water map.

In Figure 28, an example of an automatic pixel-based threshold is given. As described in Section 4.1.5, the pixel-wise threshold relates on the dynamic behaviour of each single pixel. In this case, both S-1 and S-2 water maps are used within the calculation. One can clearly distinguish between river courses and other land cover areas. Additionally it can be seen that areas where diverging properties between S-1 and S-2 maps can be observed, tend to have a lower threshold value. This circumstance is caused by the fact that the water maps of both sensor types were processed within the same time-series.

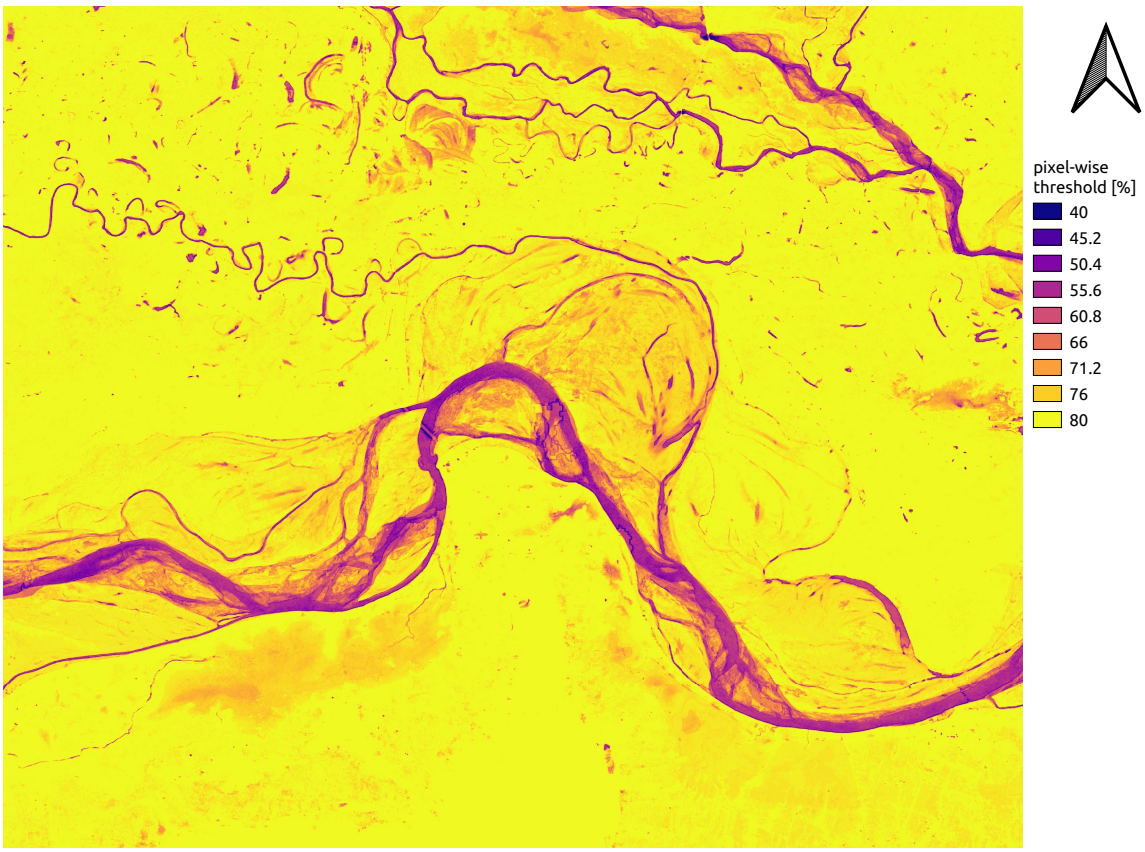
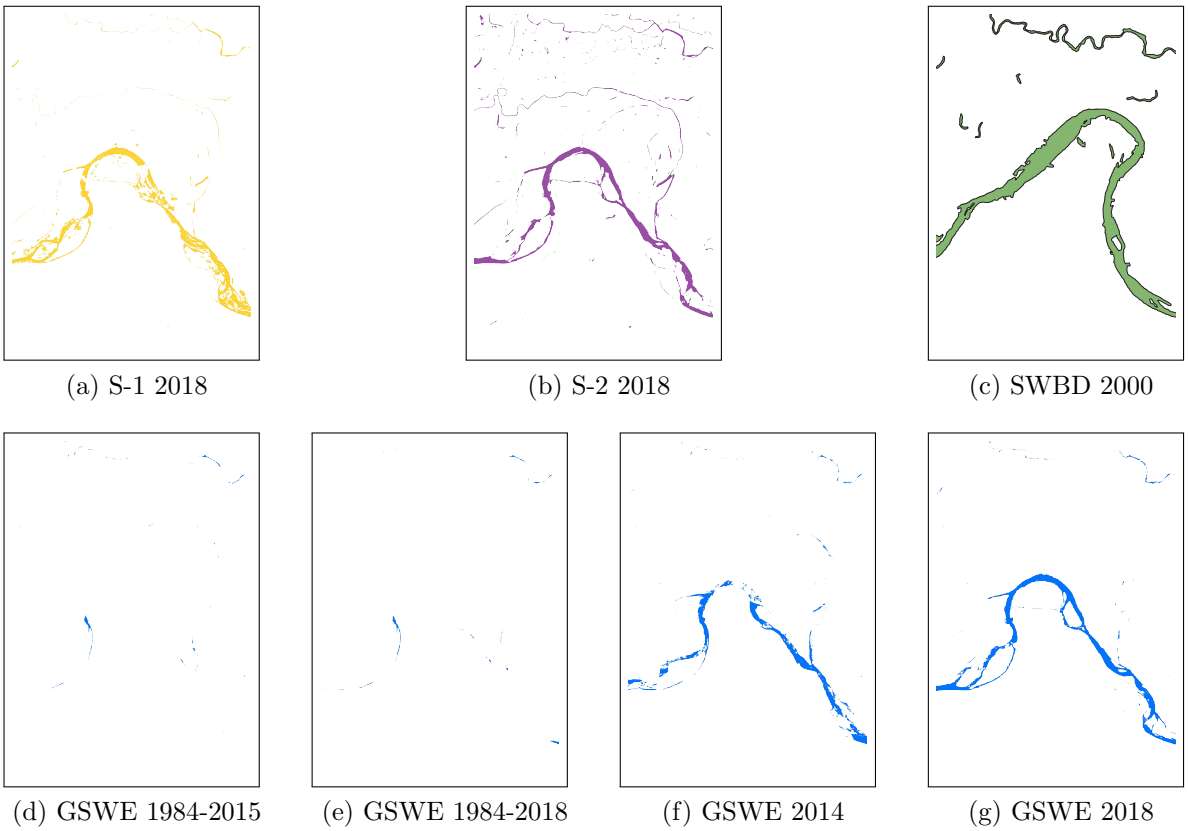


Figure 28: Overview of the pixel-wise threshold in Bihar for a combined S-1 and S-2 reference water map

As a consequence, each time a S-1 water map follows a S-2 map, the different sensor specific characteristics could yield an increase in the dynamic parameter.

5.2 External datasets comparison

Within this section, the processed reference water maps, which were generated within the scope of this thesis, are compared to representative external datasets. There are many global surface water datasets available and described in Section 2.3. In Figures 29 and 30, the water bodies from GSWE and SWBD are compared to results from S-1 and S-2 reference maps. Details on the participating surface water products can be found in Section 2.3. Reference water maps based on GSWE are obtained from the *occurrence* data of the intervals 1984-2015 and 1984-2018 and the *seasonality* data from 2014 and 2018. The main difference between these products is that the *occurrence* data was meant for the analysis of multiple years of surface water bodies, while the *seasonality* data was intended for the analysis of changes within a single year. All data from GSWE were processed by applying the same threshold like it was used for the Sentinel data. The Sentinel products were processed by using MMWC, without monthly normalisation and a pixel-wise threshold of 85 % (Bihar) and 70 % (Zaragoza).



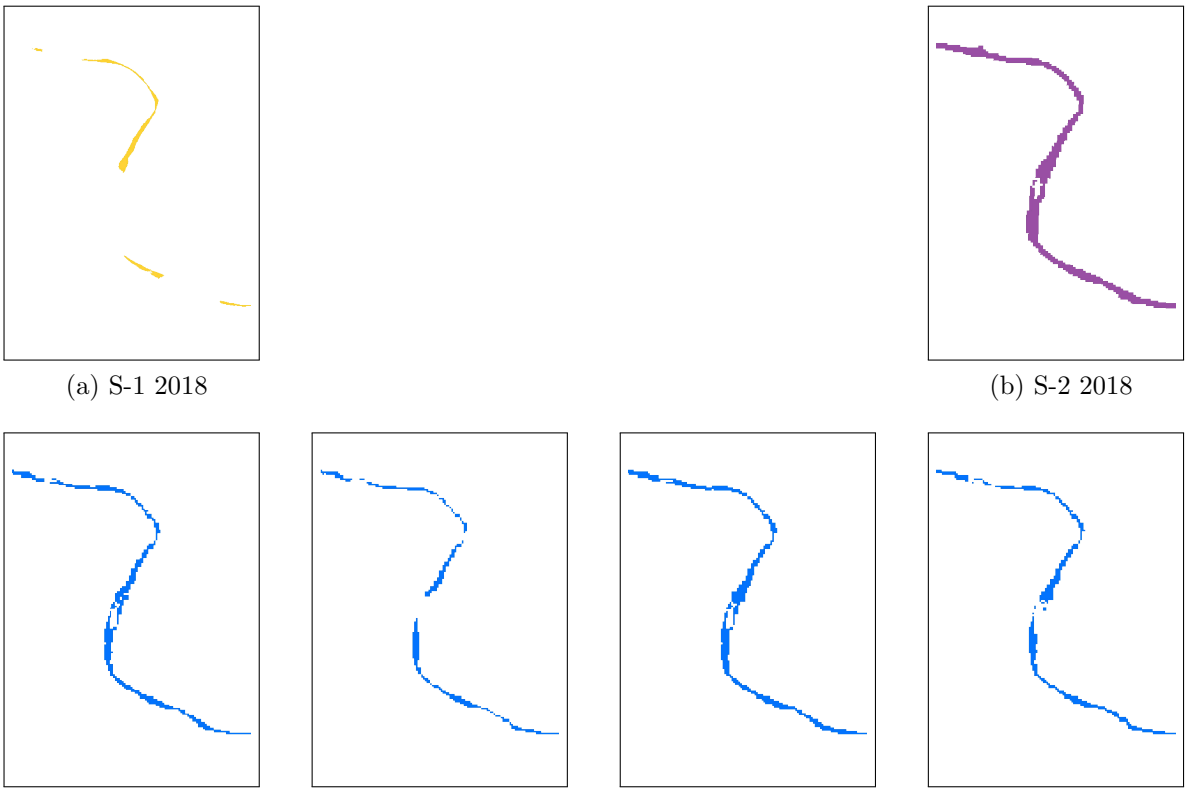
(a) S-1 2018 (b) S-2 2018 (c) SWBD 2000

(d) GSWE 1984-2015 (e) GSWE 1984-2018 (f) GSWE 2014 (g) GSWE 2018

Figure 29: Comparison of multiple surface water products in Bihar (yellow = S-1, violet = S-2, green = SWBD, blue = GSWE)

Due to its diverging observation year compared to all other datasets, SWBD (Figure 29c) shows a different location of the Ganges river. In case of flood detection tasks the resulting inundated areas would be heavily distorted. Within regions, where the water body locations are presumed to be constant, older datasets could still be relevant. Another consequence of the temporal changes of the observed water bodies are the incomplete results from long-term GSWE data (Figures 29d and 29e). Spatial variations of river branches over 30 years of optical observations lead to a low water occurrence and prevent a complete image of the local water bodies at this threshold. The threshold would have to be set under 50 % to get a valid water body representation. The small areas where water is present in the long-term GSWE reference map are located similarly in the old (Figure 29d) and the new dataset (Figure 29e). This fact might be caused by topographical constraints. Beside the already mentioned differences between S-1 (Figure 29a) and S-2 (Figure 29b) one can notice some similarities between S-2 reference map and the corresponding GSWE map (Figure 29g). For the calculation of both reference water maps, optical data (S-2 and Landsat), the same threshold (> 85 %) and the same period of time (year 2018) was used. A narrow river meander on the right side is only present in the S-2 map (Figure 29b) which could be caused by the better spatial resolution of the

original S-2 observations or some details in the processing chains. The year 2014 (Figure 29f) varies considerably from the year 2018 (Figure 29g) which supports the assumption that even a few years difference between flood event and reference map observations could cause dramatic uncertainties in the reference water map.



(a) S-1 2018 (b) S-2 2018

(c) GSWE 1984-2015 (d) GSWE 1984-2018 (e) GSWE 2014 (f) GSWE 2018

Figure 30: Comparison of multiple surface water products in Zaragoza (yellow = S-1, violet = S-2, blue = GSWE)

In Figure 30, the same type of comparison is shown for the region surrounding Zaragoza. As already mentioned, the S-1 reference water map is incomplete in case of the Elbro river in Zaragoza. This incompleteness can be seen in Figure 30a. The most continuous representation of the local water bodies is given by S-2 (Figure 30b). SWBD does not have any data available for this area of interest. Since there are water bodies covered by SWBD in eastern direction, where the river becomes wider, these local water areas are assumed to be nevertheless too narrow to be part of SWBD. In contrast to the results of Bihar, the GSWE reference maps obtained by long-term time-series (Figure 30c and 30d) deliver a more complete representation. In general all GSWE reference maps (Figures 30c - 30f) show a similar result. Only the Figure 30d stands out from the GSWE datasets. The data, which was included from the time period 2016 - 2018, seems to significantly alter the shape of the corresponding reference water map (Figure 30d). This finding is most obvious, when comparing the new GSWE long-term version

(Figure 30d) to the previous one (Figure 30c). Both effects indicate that Zaragoza is hydrologically more static than Bihar and errors within flood detection, which are caused by uncertainties of the reference water maps, are not that likely. If the water body of interest is a narrower river, the selection of the sensor has to be considered to avoid an incomplete representation.

Regarding both areas of interest, a selection of a valid reference water map has to be done with caution. One has to consider the sensor specific characteristics, the size of the local water bodies and the hydrological behaviour of the area of interest.

5.3 Flood detection

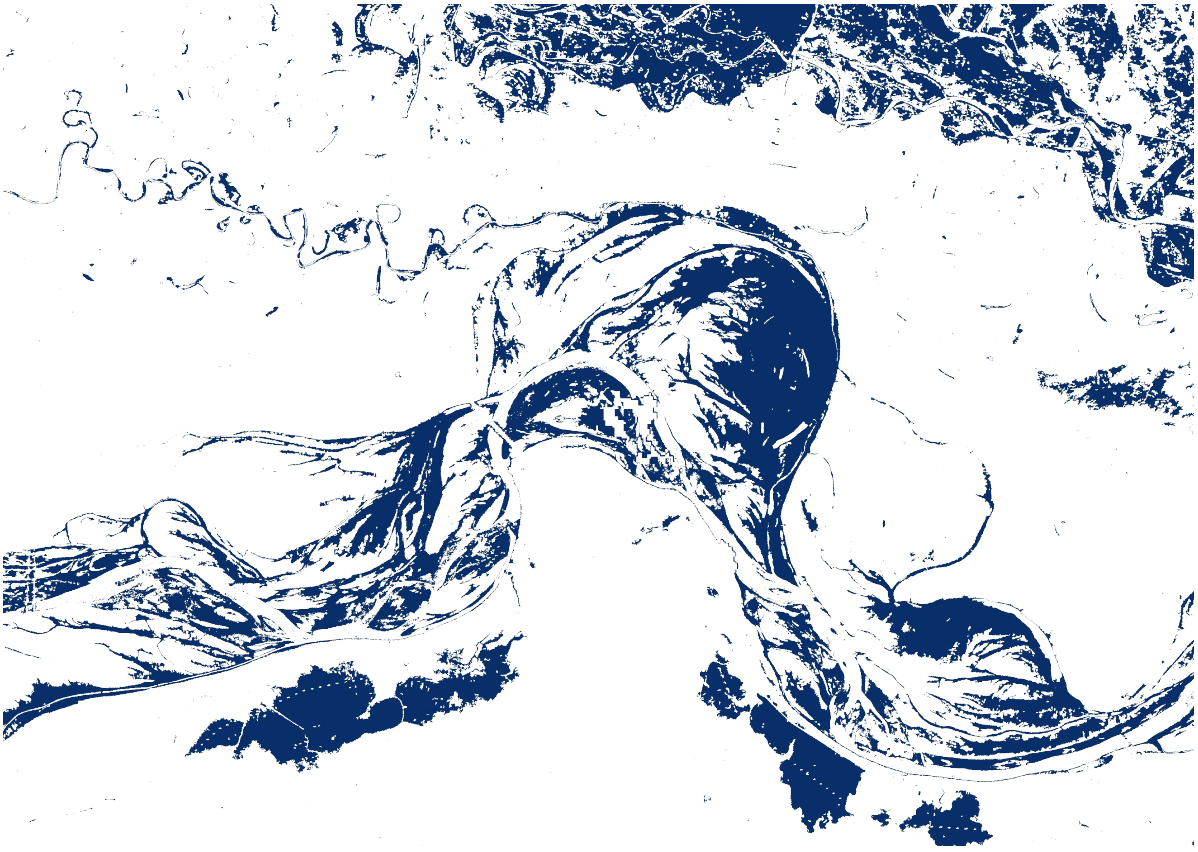


Figure 31: Flooded areas calculated by subtracting a S-1 reference water map from the S-1 water map (date: 2017-8-20)

Due to fact that clouds often prevent optical satellite sensors to deliver data during flood events, flood detection normally relies on SAR data. As a consequence, it is an obvious choice to select a S-1 reference water map to detect flooded areas. This approach was used to calculate the results (=*flooded areas*) which are shown in Figure 31. Therefore, a S-1 reference map is subtracted from a water map (=*flood map*), which was observed during a flood event at 2017-8-20.

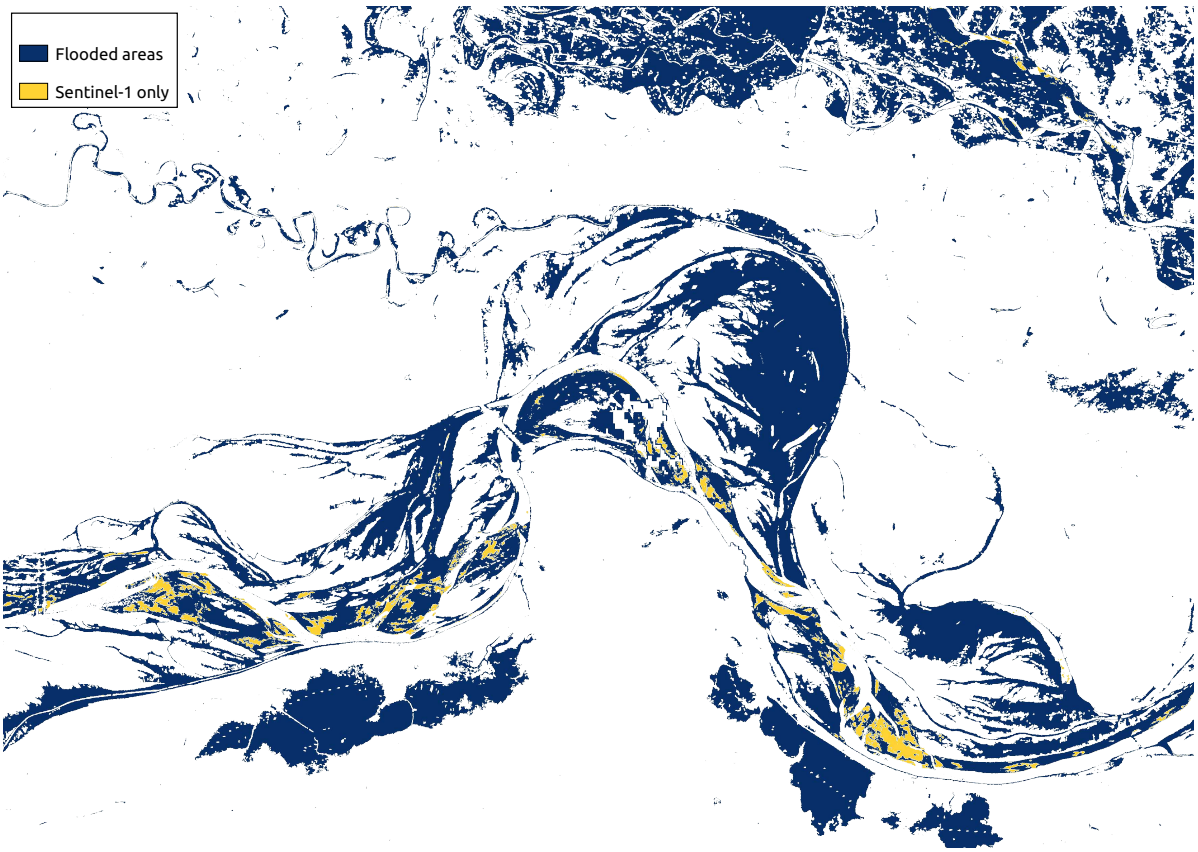


Figure 32: Comparison of the result of a S-1 based flood detection and the areas determined as overestimation

As mentioned in Section 5.1.1, one can determine the areas, where S-1 reference map shows overestimation by analysing the differences of S-1 and S-2 reference water map. Based on this information, the resulting flooded areas of Figure 31 can be reviewed. In Figure 32, an overlay of the resulting flooded areas (coloured in blue) and the areas of which are only part of the S-1 reference water map (coloured in yellow) is shown. Since these yellow coloured areas are considered to be caused by overestimation (for details see Section 5.1.1), they should not be part of the reference water map. Further on, the same areas would be classified as flooded, if they would have been removed from the reference water map. Since the discussed areas were originally part of the S-1 reference map, it can be assumed that these areas are flood-prone as they were not excluded by the S-1 processing chain's exclusion mask (for details see Section 3.2.1). By applying this correction, it could be shown that this approach is able to improve the S-1 based flood detection.

The flood detection process presented above is done by a flood map observed by S-1 and a reference water map which is processed from S-1 water maps. Figure 33 deals with the question what an effect can be expected if a S-2 reference water map is used instead. It can be seen that the yellow areas, which are discussed previously and are shown in Figure 32, are correctly defined as flooded areas. This is caused by the fact

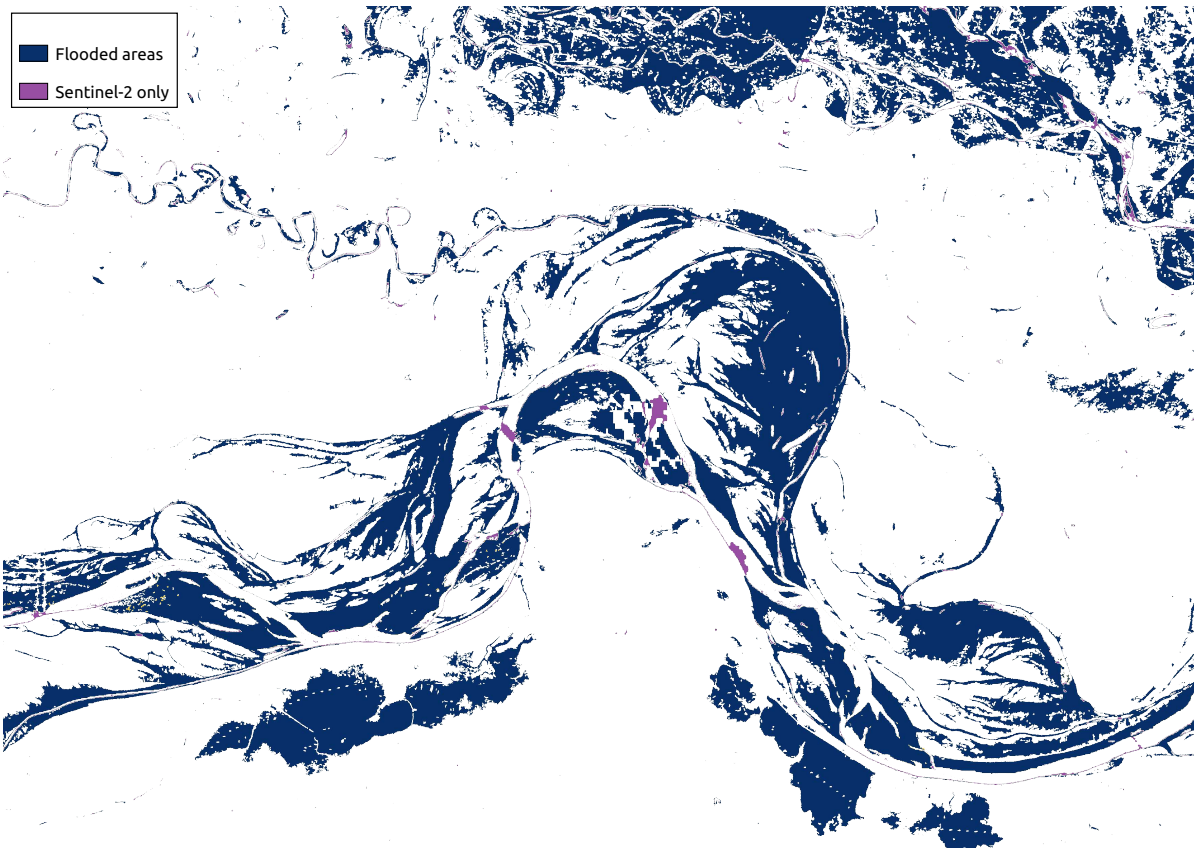


Figure 33: Result of a flood detection based on a S-1 based flood image and a S-2 reference water map. The result is overlapped by the areas which are only observed by S-1

that these areas are not part of the S-2 reference water map which was used in this flood detection approach. Additionally the areas (coloured in violet) which are excluded by the exclusion mask of the S-1 processing chain can be seen. Due to the fact that the flood image is also calculated by the S-1 processing chain there are no flooded areas intersecting the excluded areas. If there would be an intersection, the corresponding areas would have to be removed from the flooded areas. Besides of the areas, which are covered by the exclusion mask, some narrow river meanders are only observed by S-2. As mentioned in Section 5.1.1, radar backscatter could be increased in narrow parts of a river and subsequently no water is detected. In this case the impacted areas should be removed from the flooded areas as they are actually part of the reference water map. It can be shown that the diverging properties of the S-1 and S-2 reference water maps can be used to improve the flood detection process even if different sensors are used for the reference water map and the flood map.

In Section 5.1.1, difference maps are shown which display a distinction between areas, which are only observed by S-2 or S-2, and those areas, where both sensors agree. An area, which is only observed by one sensor, can either be an overestimation of the corresponding sensor or an underestimation of the other sensor.

In this Section it was described, how the knowledge of the difference between multiple sensor types can be used to improve the flood detection process. The requirement for this kind of improvement is a reasonable interpretation of the observed sensor differences which should allow the user to distinguish between over- or underestimation of a certain sensor. As a consequence, a local analysis has to be done to verify the assumptions. If an over- or underestimation is detected, the corresponding reference water mask has to be either reduced or enhanced before the flood detection process. The observed flood map could be affected by sensor specific characteristics as well. Since it is not likely that another sensor is available to detect theses potential errors, alternative methods would have to be found. One option to evaluate this kind of situation might be the exclusion mask, which is used within the S-1 processing chain (described in Section 3.2.1).

6 Conclusion and Outlook

The satellite-based detection of inundated areas during flood events (i.e. flood detection) is an essential task of Earth Observation (EO). Therefore, reference information is needed which represents the local water body's extent and location under normal conditions. By comparing the observed water extent during a flood event and the reference information, the actual flood extent can be determined. As a consequence, uncertainties within the reference information could massively influence the results of flood detection and the reference's choice should therefore get appropriate attention.

In the scope of this thesis, a flexible work-flow which delivers reference information from a time-series of classified water areas (i.e. water maps) was presented. Therefore, the percentage of water detections compared to the overall number of valid observations was defined pixel-wise and named Water Occurrence (WO). By applying a WO threshold, the resulting reference water map was calculated. Additionally, a method to automatically define the WO threshold was shown.

The main objective of this thesis was the study of differences in the resulting reference water maps caused by modifications of the presented work-flow. Therefore, parameters like the considered period of time, the used sensors or the WO threshold value can be set according to local conditions. Since there is no commonly used definition of a good reference, the optimisation is mainly based on visual interpretation.

As a result, the observed effects of different modifications of the presented work-flow should support potential users to produce reasonable reference information. In the scope of this thesis, regions where temporal changes of the surface water areas appear quite often (hydrologically dynamic) were compared with regions where the water areas are hardly affected by temporal changes (hydrologically static). The experiments were performed using data from the hydrologically dynamic region of Bihar (India) and the hydrologically more static region of Zaragoza (Spain). Data were obtained from Sentinel-1 (S-1) and Sentinel-2 (S-2) satellites in the years of 2017 and 2018.

By comparing the reference maps produced from water maps which were obtained by S-1 or S-2, it could be shown that S-1 regionally suffers from overestimation of inundated areas, mostly caused by a confusion between wet sand and actual surface water. Additionally, the radar observation of S-1 showed some underestimation in case of urban areas and narrow rivers. The estimation of the optical observation of S-2 was found to be more representative but due to the lack of information during cloud cover S-2 data are not accessible for each observation. Based on the knowledge of these sensor specific differences the result of flood detection can be improved. Water classified areas which are assumed to be caused by overestimation can be removed from the reference water map. In case of detected underestimation, the corresponding areas can be included into the reference water map.

Especially in case of hydrologically dynamic regions (like Bihar), the setting of the presented parameters (e.g. considered period of time, used sensors, WO threshold value) has a large impact on the resulting reference water map. Besides of the used sensor, the WO threshold was shown to be the major influencing factor. The automatic threshold approach could increase objectivity and decrease incompleteness. On the other hand, a

well-founded analysis and the consideration of expert knowledge can only be performed by the use of a manually predefined threshold. A WO threshold of 90 % is suggested in [19] and is recommended as a good compromise. The considered period of time defines which water maps will be included within the reference information calculation. If the considered period of time is selected too long (e.g.: several years), the resulting reference water map was shown to be affected by incompleteness and the processing time increases dramatically. On the other hand, if it is selected too short, not all seasonal variations would be considered and the error rate would be increased due to the larger influence of single water maps. As a compromise, reasonable results could be achieved by only considering those months whose water extent is close to the yearly median water extent.

In case of further research, additional areas of interest which feature other challenges for surface water detection should be considered. These are for example forests, polar landscape or deserts. Further findings on the diverging characteristics of S-1 and S-2 could be obtained and applied for an improved flood detection. Therefore, the definition of a resulting reference water map should be dependent on the geographic region, hydrological behaviour and land cover to better employ each sensor. In order to increase the amount of automation and objectivity, a statistical definition process of the WO threshold should be developed.

List of Figures

List of Tables

References

[1] Bing maps tile service. https://docs.microsoft.com/en-us/bingmaps/rest-services/directly-accessing-the-bing-maps-tiles. Accessed: 2019-9-24.

[2] Copernicus open access hub. https://scihub.copernicus.eu/dhus/#/home. Accessed: 2019-9-24.

[3] Dartmouth flood observatory. https://disasters.nasa.gov/programs/dartmouth-flood-observatory. Accessed: 2019-08-29.

[4] Dlr -earth observation center - team: Naturgefahren. https://www.dlr.de/eoc/desktopdefault.aspx/tabid-6429/10561_read-36329/. Accessed: 2019-9-25.

[5] Emsr279: Flood in the ebro river basin, spain. https://emergency.copernicus.eu/mapping/list-of-components/EMSR279. Accessed: 2019-08-18.

[6] European flood awareness system. https://efas.eu/. Accessed: 2019-08-29.

[7] European union (eu) floods directive 2011. https://www.flussgebiete.nrw.de/system/files/atoms/files/dok_hwrmrl_2011_02_16_reporting_sheets_en.pdf. Accessed: 2020-1-2.

[8] Gdal documentation. https://gdal.org/index.html. Accessed: 2019-08-19.

[9] Global flood awareness system. http://www.globalfloods.eu/. Accessed: 2019-08-29.

[10] Planet - daily satellite imagery and insights. https://www.planet.com/. retrieved data from Planet Online plattform.

[11] Qgis project. https://qgis.org/en/site/index.html. Accessed: 2019-08-19.

[12] Sentinel-2 - about the launch. http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/About_the_launch. Accessed: 2019-08-18.

[13] Srtm water body data. https://dds.cr.usgs.gov/srtm/version2_1/SWBD/SWBD_Documentation/SWDB_Product_Specific_Guidance.pdf. Accessed: 2019-04-11.

[14] Un office for disaster risk reduction. https://www.unisdr.org/. Accessed: 2019-09-23.

[15] Web mapping service. https://www.opengeospatial.org/standards/wms. Accessed: 2019-9-24.

[16] Worldclim version2. http://worldclim.org/version2. Accessed: 2019-08-19.

[17] Zaragoza. https://www.britannica.com/place/Zaragoza-Spain. Accessed: 2019-08-13.

[18] Filipe Aires, Leo Miolane, Catherine Prigent, Binh Pham, Etienne Fluet-Chouinard, Bernhard Lehner, and Fabrice Papa. A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *Journal of Hydrometeorology*, 18(5):1305–1325, 2017.

[19] Filipe Aires, Catherine Prigent, Etienne Fluet-Chouinard, Dai Yamazaki, Fabrice Papa, and Bernhard Lehner. Comparison of visible and multi-satellite global inundation datasets at high-spatial resolution. *Remote sensing of environment*, 216:427–441, 2018.

[20] Jörg Albertz. *Einführung in die Fernerkundung: Grundlagen der Interpretation von Luft-und Satellitenbildern*. Wiss. Buchges., 2009.

[21] Luc Bertels, Bruno Smets, and Davy Wolfs. Dynamic water surface detection algorithm applied on proba-v multispectral data. *Remote Sensing*, 8(12):1010, 2016.

[22] Surekha Borra, Rohit Thanki, and Nilanjan Dey. *Satellite Image Analysis: Clustering and Classification*. Springer, 2019.

[23] Mark L Carroll, John R Townshend, Charlene M DiMiceli, Praveen Noojipady, and Robert A Sohlberg. A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4):291–308, 2009.

[24] P Dayal, Allen G Noble, and Ashok K Dutt. Bihar. https://www.britannica.com/place/Bihar. Accessed: 2019-08-13.

[25] Matthias Drusch, Umberto Del Bello, Sébastien Carlier, Olivier Colin, Veronica Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, et al. Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote sensing of Environment*, 120:25–36, 2012.

[26] Min Feng, Joseph O Sexton, Saurabh Channan, and John R Townshend. A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, 9(2):113–133, 2016.

[27] Etienne Fluet-Chouinard, Bernhard Lehner, Lisa-Maria Rebelo, Fabrice Papa, and Stephen K Hamilton. 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, 2015.

[28] Matthew C Hansen, Peter V Potapov, Rebecca Moore, Matt Hancher, SAA Turubanova, Alexandra Tyukavina, David Thau, SV Stehman, SJ Goetz, Thomas R Loveland, et al. High-resolution global maps of 21st-century forest cover change. *science*, 342(6160):850–853, 2013.

[29] Renaud Hostache, Patrick Matgen, and Wolfgang Wagner. Change detection approaches for flood extent mapping: How to select the most adequate reference image from online archives? *International Journal of Applied Earth Observation and Geoinformation*, 19:205–213, 2012.

[30] Chang Huang, Ba Duy Nguyen, Shiqiang Zhang, Senmao Cao, and Wolfgang Wagner. A comparison of terrain indices toward their ability in assisting surface water mapping from sentinel-1 data. *ISPRS International Journal of Geo-Information*, 6(5):140, 2017.

[31] Karl Kraus and Werner Schneider. *Fernerkundung. 1. Physikalische Grundlagen und Aufnahmetechniken: mit 15 Tabellen*. Dümmler, 1988.

[32] Claudia Kuenzer, Huadong Guo, Juliane Huth, Patrick Leinenkugel, Xinwu Li, and Stefan Dech. Flood mapping and flood dynamics of the mekong delta: Envisat-asar-wsm based time series analyses. *Remote Sensing*, 5(2):687–715, 2013.

[33] Céline Lamarche, Maurizio Santoro, Sophie Bontemps, Raphaël d'Andrimont, Julien Radoux, Laura Giustarini, Carsten Brockmann, Jan Wevers, Pierre Defourny, and Olivier Arino. Compilation and validation of sar and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. *Remote Sensing*, 9(1):36, 2017.

[34] Joel Lawhead. *Learning geospatial analysis with Python*. Packt Publishing Ltd, 2015.

[35] Bernhard Lehner, Kristine Verdin, and Andy Jarvis. New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10):93–94, 2008.

[36] Yu Li, Sandro Martinis, Simon Plank, and Ralf Ludwig. An automatic change detection approach for rapid flood mapping in sentinel-1 sar data. *International journal of applied earth observation and geoinformation*, 73:123–135, 2018.

[37] Carsten Pathe, Wolfgang Wagner, Daniel Sabel, Marcela Doubkova, and Jeffrey B Basara. Using envisat asar global mode data for surface soil moisture retrieval over oklahoma, usa. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2):468–480, 2009.

[38] Jean-François Pekel, Andrew Cottam, Noel Gorelick, and Alan S Belward. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418, 2016.

[39] Catherine Prigent, Elaine Matthews, Filipe Aires, and William B Rossow. Remote sensing of global wetland dynamics with multiple satellite data sets. *Geophysical Research Letters*, 28(24):4631–4634, 2001.

[40] Pinzeng Rao, Weiguo Jiang, Yukun Hou, Zheng Chen, and Kai Jia. Dynamic change analysis of surface water in the yangtze river basin based on modis products. *Remote Sensing*, 10(7):1025, 2018.

[41] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[42] Maurizio Santoro, Urs Wegmüller, Céline Lamarche, Sophie Bontemps, Pierre Defourny, and Olivier Arino. Strengths and weaknesses of multi-year envisat asar backscatter measurements to map permanent open water bodies at global scale. *Remote Sensing of Environment*, 171:185–201, 2015.

[43] Stefan Schlaffer, Patrick Matgen, Markus Hollaus, and Wolfgang Wagner. Flood detection from multi-temporal sar data using harmonic analysis and change detection. *International Journal of Applied Earth Observation and Geoinformation*, 38:15–24, 2015.

[44] André Twele, Wenxi Cao, Simon Plank, and Sandro Martinis. Sentinel-1-based flood mapping: a fully automated processing chain. *International Journal of Remote Sensing*, 37(13):2990–3004, 2016.

[45] Jan Verbesselt, Achim Zeileis, and Martin Herold. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, 123:98–108, 2012.

[46] Charles Verpoorter, Tiit Kutser, David A Seekell, and Lars J Tranvik. A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18):6396–6402, 2014.

[47] Marc Wieland, Yu Li, and Sandro Martinis. Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network. *Remote Sensing of Environment*, 230:111203, 2019.

[48] Iain H Woodhouse. *Introduction to microwave remote sensing*. CRC press, 2017.

[49] Hanqiu Xu. Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, 27(14):3025–3033, 2006.

[50] Dai Yamazaki, Mark A Trigg, and Daiki Ikeshima. Development of a global˜ 90 m water body map using multi-temporal landsat images. *Remote Sensing of Environment*, 171:337–351, 2015.