SAR-based Flood Monitoring

Comparison of Synthetic Aperture Radar sources on the Water-Land Boundary estimation for flood events in the Netherlands

Femke M. Bulsing

Delft University of Technology

TU Delft

SENSAR

Rijkswaterstaat
Ministerie van Infrastructuur en Waterstaat

SAR-based Flood Monitoring

Comparison of Synthetic Aperture Radar sources on the Water-Land Boundary estimation for flood events in the Netherlands

by

Femke M. Bulsing

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended on Wednesday February 26, 2025 at 15.45.

Student Number:	5415047	
Project Duration:	May, 2023 - February, 2025	
Faculty:	Civil Engineering and Geosciences	
Thesis committee:	Reinier Oost (Daily supervisor)	Sensar
	Dr. ir. Paco Lòpez-Dekker (Chair)	TU Delft, Geoscience and Remote Sensing
	Dr. Roderik Lindenbergh	TU Delft, Geoscience and Remote Sensing
	Prof. dr. ir. Nick van de Giesen	TU Delft, Water Management

Confidential information of the thesis is included in an appendix that is left unpublished.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Preface

After a long journey, I can finally write this preface. My master's journey began in 2020, during the lockdown. Despite the challenges, my passion for earth sciences kept me going. Not just the curiosity about how the Earth works and but also learning how to measure its processes. As a child, I wondered why the grass was green and how waves could become so wild in the ocean, and it's that same fascination with Earth's complexity that continues to inspire me in this field every day. This thesis aims to help measure this complexity, which is especially important as climate change presents increasing global challenges that we will all have to face someday.

Working with passionate people, like my supervisors, has shown me that this field is about more than data. It's about helping our community in ways that may not be possible yet, but could be possible in the future. Thanks to Reinier Oost for our many educational sessions and showing me the importance of communication in data. Paco López-Dekker for his valuable feedback on this topic and enduring sessions on the quality of the data. Roderik Lindenbergh and Nick van de Giesen for their guidance, and Rinus Schroevers for his helpful and honest feedback on the impact of the research. Not to forget, Marian Roodenburg who also supported me a lot with everything around the thesis.

Graduating has been a major goal. Whenever I felt doubtful, my colleagues, friends, family, and partner were always there for me, and I am so grateful for all their support. Let this be a reminder that achieving your goals often requires the support of those who believe in you. Surround yourself with people who motivate and uplift you, it makes all the difference.

Many people have contributed to this journey, and much effort has been invested. With that, I invite you to read the remainder of the thesis. May it inspire you to address the challenges we face and be encouraged by the technology that helps us overcome them.

Femke M. Bulsing
Delft, February 2025

Abstract

Floods are natural hazards with severe impacts, and their frequency and intensity are increasing due to climate change. Synthetic Aperture Radar (SAR) satellites are widely used for flood mapping, as they operate independently of weather and time of day. This thesis examines the potential of SAR-based flood monitoring through a case study of the July 2021 flood in Limburg, the Netherlands. Comparing Capella Space on-demand X-Band imagery (sub-meter resolution) with Sentinel-1 open-source C-Band imagery (5 m x 20 m resolution). The Water-Land Boundary was determined by estimating flood extent and water levels, with multiple methods evaluated for the distinct products from both data sources. Capella Space provided a single image acquisition at the flood's peak, with a thresholding method used to classify flooded pixels. For Sentinel-1, an Amplitude Time Series Analysis (ATSA) was applied to data from 2017 to 2024 to identify flood-related outliers. The water levels are estimated from flood extent edges with the national LiDAR DEM (AHN4).

Evaluation of the modeled results using an error matrix at the acquisition time showed that 67% and 68% of the pixels were correctly classified from the flood extents derived from Capella Space and Sentinel-1, respectively. The maximum flood extent from Sentinel-1 data decreased to 45% correct classification when compared to the modeled results at the peak of the flood. This is consistent with the acquisition times, which were taken before and two days after the flood peak, missing the peak flood moment. SAR-based water levels showed an overall precision of 0.141 m for Capella Space and 0.156 m for Sentinel-1. Agreement with water level gauge measurements was better in flatter, less vegetated areas and lower in steep, vegetated areas. Achieving consistent 20 cm water level accuracy (as required by the Dutch Ministry of Infrastructure and Water Management), across the study area remains complex. Both methods are prone to false positives and negatives, especially in areas with steep slopes, narrow canals, high vegetation, or roads. False classifications result in inaccurate flood extents, thus decreasing water level accuracy. Higher resolution of Capella Space images provided better alignment with the maximum flood extent, while Sentinel-1 images have wider coverage but missed the timing of the flood peak. In the end, the choice of SAR-system depends on timing, surface characteristics, and mapping extent needs.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
AOI	Area of Interest
AHN	Actueel Hoogtebestand Nederland
ATSA	Amplitude Time Series Analysis
CET	Central European Time
CEST	Central European Summer Time
dB	Decibel
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
EMR	ElectroMagnetic Radiation
ENW	ExpertiseNetwerk Waterveiligheid
FN	False Negatives
FP	False Positives
GRD	Ground Range Detected
HH	Horizontal Horizontal polarization
InSAR	Interferometric Synthetic Aperture Radar
IPS	Intermittent Persistent Scatterers
IQR	InterQuartile-Range
LMW	Landelijk Meetnet Water
LiDAR	Light Detection and Ranging
MERLIN	coMplex sElf-supeRvised despeckLINg
NAP	Normaal Amsterdams Peil
PS	Persistent Scatterers
RADAR	RAdio Detection And Ranging
RoIs	Regions of Interest
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SLC	Single Look Image
TAW	Tweede Algemene Waterpassing
TN	True Negatives
TP	True Positives
TSA	Time Series Analysis
VH	Vertical Horizontal polarisation
VHR	Very High Resolution
VV	Vertical Vertical Polarisation
UTC	Universal Time Coordinated

1

Introduction

1.1. Background

Floods are one of the natural disasters that will impact the world more frequently as climate change intensifies, leading to destruction of property, agriculture, and, in severe cases, causing casualties among both humans and animals [1]. Flooding is a phenomenon where land, that is supposed to be dry, becomes inundated. Different categories can be made depending on the processes, place and timing: pluvial floods, river floods, groundwater floods, coastal floods [2]. The different types are explained below [3] [4]:

- **Pluvial Floods**: Pluvial flooding occurs due to extreme rainfall events. As a result, the amount of water exceeds the capacity of the land surface. This can take place independently of overflowing permanent water bodies. Pluvial flooding could occur in any urban or rural location and at any distance from existing water bodies. Pluvial floods can be divided into surface-water floods and flash floods. Surface-water floods are most common in urban areas and can be related to exceeding capacities of local drainage systems. Flood water spreads onto streets and structures that are located nearby overflowing drainage systems. This type of flood is generally shallow, with water levels of less than 1 meter deep. Flash floods are related to intense and high-velocity torrents of water that occur in terrains with high elevation differences. They are characterized by large amounts of water falling down elevated terrain in a short amount of time. A potential cause of flash floods can be sudden releases from upstream dikes or dams. Flash floods can cause catastrophic damage to places and people living downstream.
- **River Floods**: River Flooding, also referred to as Fluvial Flooding, is related to rising water levels in permanent water bodies such as rivers, lakes, or streams. Due to heavy rainfall or snow, water levels exceed the elevation of the banks and water overflows onto the land. River flooding can be simulated by flood prediction models, including: forecasts of precipitation, measurements of water levels, and conditions of the surrounding land (such as soil saturation capacities). The damage of river flooding is often related to the duration and intensity of the rising water levels.
- **Groundwater Floods**: Groundwater flooding is caused by rising water tables that have been stored in the ground, leading to land inundation. This type of flood can be caused by long and heavy rainfall that ultimately exceeds the capacity of the ground reservoir. Groundwater floods are usually local floods that occur slowly over time, but can lead to severe damage if they continue for an extended period.
- **Coastal Floods**: Coastal flooding occurs when sea levels along coastlines or estuaries rise and overflow onto the surrounding land. The magnitude in which sea water can penetrate inland is dependent on the topography of the land and the strengths of its coastal defenses. Coastal floods are related to high tides or storm surge. Storm surges can be caused by low-pressure weather systems such as tropical and post-tropical cyclones.

One of the main drivers for flooding can be extreme precipitation. Due to large increases in precipi-

tation, water exceeds the capacity of the natural or artificial drainage system or causes a rise in sea level. Some variables such as snow-water equivalent, stream morphology and river or catchment engineering play a large role in the severity of the flood. The understanding goes deeper as the amount of research and studies dedicated to floods increases. However, as flooding is a complex mechanism that includes weather, hydrology and hydrodynamics, the prediction of flood parameters remains difficult for scientists.

In the Netherlands, the risk of coastal flooding is significant, as a large part of the country lies below sea level. Further inland, the main indications for floods are extreme precipitation events. Despite the implementation of the Delta Works in the Netherlands, 59% of the country's land area remains at risk of flooding. This includes: 26% located below sea level, 29% vulnerable to riverbank overflow, and 4% located on the seaward or riverside of the dikes [5].

As climate change intensifies, the intensity of extreme precipitation events is also expected to increase. Based on multi-model simulations from global coupled climate models, changes in total precipitation intensity are depicted in Figure 1.1. The intensity of precipitation in Figure 1.1 is defined as the total annual precipitation divided by the number of wet days. As can be seen in the figure, the risk of floods in coastal, urban and other low-elevated regions will become higher in the future [2] [6]. In the Netherlands, the risks of floods due to rising sea levels are among the highest in Europe, particularly in terms of the population at risk. The map in Figure 1.2 shows the amount of expected sea level rise per country for the population count in 2020. It shows the risk of a 100-year coastal flood event and the risk of the population under different climate scenarios. These figures show that nearly 1 million people in the Netherlands are at risk of a 2-meter rise in sea level. In addition to the danger to people's safety, there are also social, environmental and economic impacts that can become extremely high [2].

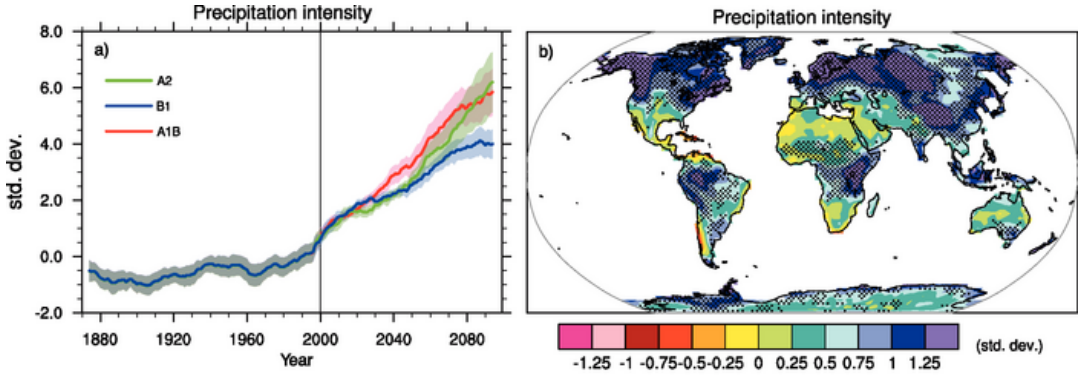


Figure 1.1: Changes in weather extremes based on multi-model simulations computed from nine global coupled climate models, results are adapted from Tebaldi et al. (2006). (a) The globally averaged changes of the precipitation intensity computed for the low (SRES B1), middle (SRES A1B) and high (SRES A2) scenario. (b) Spatial pattern of the changes in simulated precipitation intensity between two 20-year means simulated for 2080-2099 and 1980-1999 [2].

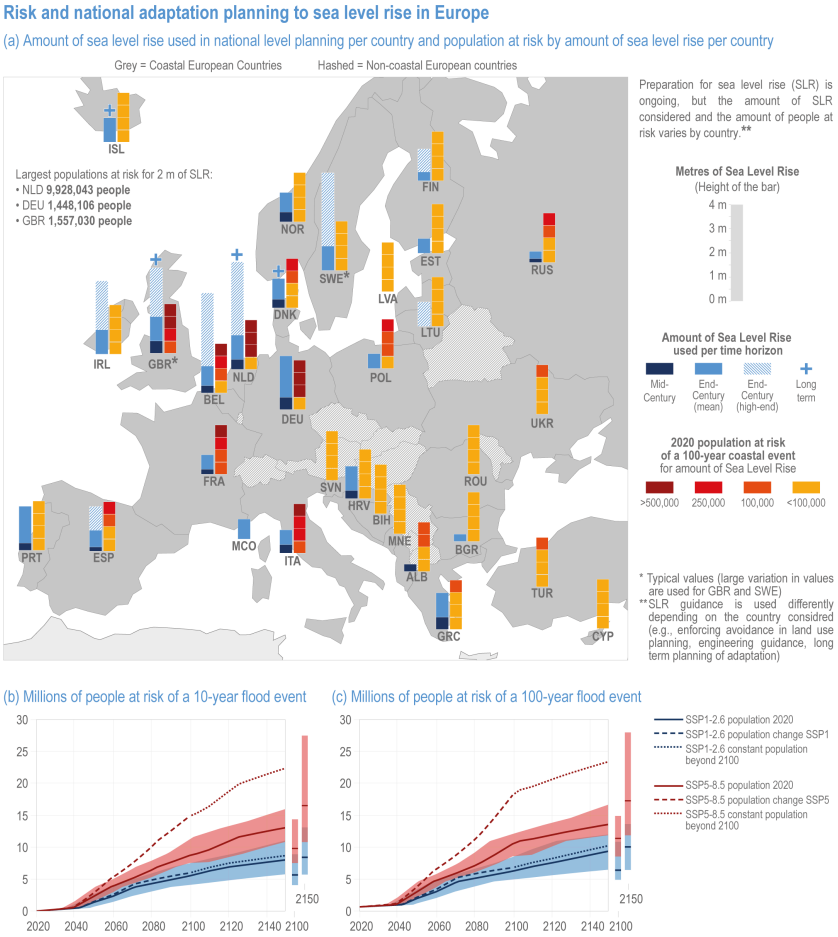


Figure 1.2: The vulnerability and national planning in Europe based on predicted Sea Level Rise (SLR) of different climate scenario's [2].

In July 2021, large parts of Belgium, Germany and the Netherlands were hit by extreme amounts of rainfall and floods. The precipitation caught in the Meuse and Rhine catchments led to excessive water levels reaching outside the river banks and inside the land. The damage was high in the entire area, not to mention the hundreds of casualties reported by the impact of the floods in Belgium and Germany. Under normal conditions, daily precipitation of more than 50 mm is considered a "heavy precipitation day" in the Netherlands [7]. The situation in 2021 was exceptional, as precipitation accumulated from 160 to 180 mm in two days. These numbers are rare in this area, especially considering the time of year. During the summer period, such numbers have not been previously recorded. Peak discharges in the Meuse catchment near Eijsden, along with regional streams, represented the highest water levels ever documented. The weather forecasts predicted large amounts of rainfall and, accordingly, adjusted the forecasts of peak flows. However, due to the rarity of the event, it was not possible to reproduce the peak flows before including the 2021 measurements.

The excess of water also led to erosion at multiple locations. The erosion pits along the bed of the 'Grensmaas' reached depths of more than 3 meters, and large amounts of sand were deposited on the riverbank. Furthermore, 95% of the plastics typically transported throughout the year were now transported within a span of only two days. This demonstrates that a large amount of debris was transported through the affected area. The flood resulted in substantial physical and economic damage and losses, significantly greater compared to the floods in 1993 and 1995. Exceptionally, most of the affected area was located within the floodplain of the regional rivers, particularly the Geul floodplain. Most of the damage and losses related to previous floods were, contrarily, found in the Meuse floodplain [8]. Analyzing all previous flood events is of great importance for further developing defense strategies of the country against natural hazards.

Figure 1.3: An overview photograph of Limburg, the Netherlands, showing a region that was affected by the flood event in 2021.

Comprehensive fact finding studies and analyses were conducted after the flood event, commissioned by the Dutch Expertise Network for Flood Protection (ENW). The rarity of the event made it difficult for the models to predict the peak flows. As mentioned above, peak flows in the Meuse near Eijsden and regional streams reached historically high water levels. However, water levels in the Meuse downstream of Roermond were lower than anticipated and did not exceed those recorded during previous peak events. This latter effect is due to the Meuse Works river program that controls the flow and flattening of the flood wave that occurs downstream of the river system [8]. hydrodynamic models are capable of simulating the spatial extent of floodwaters with a high accuracy. The timing however, of inflow and outflow of water is not yet in alignment with survey measurements in the field. Timing inaccuracies are partially related to the incorporation of winter vegetation in the models instead of summer vegetation. The type of vegetation can have an effect on the degree of saturation and roughness of the surface, that ultimately determines the flow rate. Water levels gauges from governmental institutions are located over the rivers and streams. Gauges from Rijkswaterstaat, or the Landelijk Meetnet Water (LMW), monitor water levels of the Meuse River during flood event. However, some of the gauges operated by the Water Boards along the streams, failed to maintain measuring once the water levels of the streams reached outside the banks. During the flood event, the Waterschapshuis, Rijkswaterstaat and several companies worked together on acquiring aerial-images and LiDAR data over the area affected by the flood (see Figure 1.4). During this crisis, they were required to establish a flight route in a matter of hours. One of the challenges for aircraft and optical imagery is the impact of weather conditions and cloud cover. Weather conditions are required to be safe and aircraft usually operate below the clouds.

Figure 1.4: A screenshot from a slider figure displaying the differences between two aerial photographs taken from the city of Bunde during the flooding event (left) and in a normal situation (right) [9].

As spaceborne measurement techniques are becoming more advanced, state of the art methods have made it possible to use satellites for flood mapping. These flood maps can be used either during or after a flood event. During a flood event, it can provide valuable information for evacuation services and risk management. Flood maps can help identifying the flooded and non-flooded regions and contribute to the flood extent or timing depicted by hydrodynamic models. In the period following the event, flood maps can support calibration and validation of flood extents predicted by models, as they represent the inundated conditions at a specific point in time. In the present day, several emergency services are in place that include the use of remote sensing. One of the most widely known is the Copernicus Emergency Management Service (EMS). EMS provides information on emergency responses related to natural or man-made disasters, among others floods. The product consists of Mapping Services and Early Warning Systems based on Sentinel-1 imagery [10]. The potential use of satellite imagery is acknowledged by scientists and water management authorities in the Netherlands, but implementations are limited. Results from satellite imagery often consist of either over- or underestimation of the flood extent. The possibilities of the latest sub-meter resolution satellites are expected to provide more accurate flood maps. Commercial companies that acquire the data have created complete processing chains for flood emergencies. However, as the technology is still emerging, it has not yet been fully adopted by all institutions. The products still require evaluation to assure the necessary requirements.

Satellite imagery has the advantage of wide coverage per acquisition, sources such as Sentinel-1 map the entire country within a few days. Using radar satellites during flood events is of particular interest as they can operate at any time of day and independently of weather conditions. Weather limitations are therefore relatively low compared to optical imagery. Optical satellites require the Sun's radiation that reflects back from the surface and atmosphere. It measures in the visible or near-visible part of the electromagnetic spectrum. The presence of clouds can block the view of the surface, thereby obstruct optical satellite imaging systems of capturing surface information.

Synthetic Aperture Radar (SAR) is a type of radar that is widely explored for flood mapping. SAR is an active remote sensing technique that sends out and receives radar waves. As active sensors send their own radiation, they are independent on sunlight. Radar waves are in the radiowave part of the electromagnetic spectrum (EM) and therefore capable to penetrate clouds. Figure 1.5 shows the Atmospheric Window of the Earth. This diagram shows how waves pass or get blocked by the atmosphere to the Earth's surface. Among others, molecules such as ozone, water, and carbon dioxide are present in the atmosphere. Only specific parts of the electromagnetic spectrum can pass through the atmosphere, this is called the Atmospheric Window. As can be seen from the figure, radiowaves can pass through the atmospheric window.

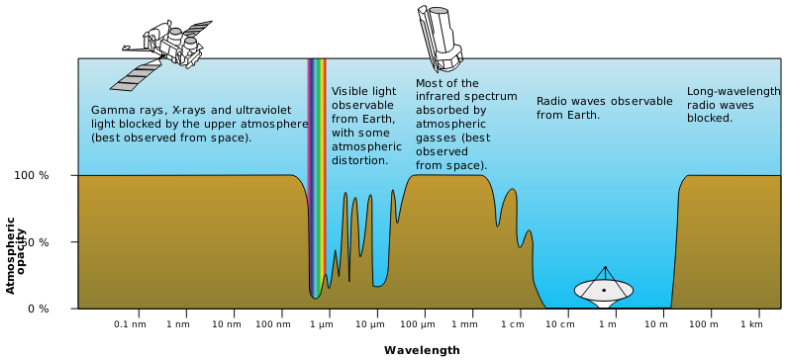


Figure 1.5: The opacity of the atmosphere related to the electromagnetic spectrum [11].

Considering that flood events are inherently related to precipitation events and dense cloud covers, SAR-derived flood maps could theoretically be a reliant source for flood mapping. The challenges of using SAR-satellites for flood events in the Netherlands are acquisition timing and spatial resolution. Floods in the Netherlands are usually relatively small in extent and of short duration, when compared to neighboring countries such as Germany. Dutch flood events usually take place in a timespan of a few

hours or days. The alignment of the satellite's acquisition to the peak of the flood is therefore difficult to accomplish. Additionally, the spatial resolution of SAR-satellites varies from sub-meter to tens of meters. For smaller streams around 20 meters, extracting the flood extent might not be feasible. Despite these challenges, the capabilities of SAR are promising and show increasing potential for usage by authorities and researchers.

1.2. Research objective

The aim of this thesis is to investigate whether SAR-based flood monitoring is accurate enough for researchers or authorities to utilize during flood events in the Netherlands. Currently, there are other resources already employed in these scenario's such as: hydrological & hydrodynamic models, aerial imagery and LiDAR imagery. Potentially, SAR-based flood maps can provide complementary data to these current resources. The opportunities of SAR can be found in different stages of a flood: evacuation management throughout the flood event, damage and insurance analysis post-flood, and calibration of hydrological & hydrodynamic models. The latter can be used for post-flood evaluation and is of significant importance for forecasting future floods. The main research question of this thesis is:

What is the value of SAR-based flood mapping as a complement to flood models, in-situ measurements and other remote sensing data for flood events in the Netherlands?

With SAR-based flood mapping, the Water-Land boundary can be derived. This can be divided into mapping of the flood extent and the water levels. Together, it describes the area (x,y) and height (z) of the surface water during the flood. With these variables the boundary between the water and land can be estimated. Current models can already compute the spatial extent of the flood with high accuracy. The complexities lie in correctly timing of water inflow and outflow during the event, that is partly related to the incorporation of winter vegetation instead of summer vegetation. Gauges can measure water levels with high accuracy, but are positioned locally along rivers and streams. Some gauges are also susceptible to failure when water reaches outside the banks. Remote sensing techniques, such as aerial images, can be used to map the affected area during the flood. However, planning of the flight plan and ensuring to fly along the peak of the flood has shown to be difficult. Additionally, aerial imagery is limited to the requirements of safe weather conditions and flying below cloud ceiling. Overall, the most important knowledge gaps in this field include the lack of data across the entire area and the challenge of acquiring images at any given moment.
Contemporary SAR companies providing high-resolution imagery, such as Iceye or Capella Space, operate mostly on-demand system, acquiring images over specific Regions of Interest (RoIs) with a lead time in the range of hours. This opens the possibilities for usage of the data in near real-time during evacuations and calibrating the timing of models. A limitation of images from on-demand companies is that their coverage area is typically limited to a few kilometers (Capella Space Sliding Spotlight: 5 km x 10 km) [12]. Open data sources of SAR provide coverage of the entire Netherlands within a few days and offer data archives starting from 2015 (ESA, Sentinel-1) [13]. Both sources contain important aspects and opportunities for knowledge gaps in this field.

To further specify the main research question, the following sub-research questions have been formulated:

- *What are the most important components of the SAR signal that provide information about the flooded state of a pixel?*
  A SAR signal contains multiple components like amplitude, phase and polarization. To delineate the flooded area, relevant components of the SAR signal should be considered for the research. By researching the different components it should become clear what significantly changes or provides valuable information about the (flooded) state of a pixel.

- *How can SAR imagery be combined with a Digital Elevation Model to accurately extract the water levels?*
  To estimate the water levels, a Digital Elevation Model is necessary to combine with flooded pixels derived from SAR imagery. This requires alignment of the two data types and the correlation between flooded pixels and the height value of the DEM.

- *How do the Water-Land Boundary estimation from on-demand sub-meter spatial resolution satellites compare to meter spatial resolution satellites with longer data archives?*
  Sub-meter resolution imagery contains more information about the area spatially. However, meter resolution imagery often comes with longer data archives and more components to the signal. It is important to find out how the information from both data sources differs. Ultimately, the decision of data source is dependent on the capabilities of the measurement technique.

- *How do the Water-Land Boundary estimations from SAR imagery compare to current resources?*
  For authorities and researchers from other fields it is important to know what to expect from SAR-based Water-Land Boundary estimations. SAR is currently not widely considered as a data source for flood events in the Netherlands. Analyzing the similarities and differences of the results is crucial to understanding the complementarity of the data. Rijkswaterstaat has established an accuracy of 20 cm for water levels, this is required for data utilization.

1.3. Method

Firstly, extensive literature research is required on current methods for SAR-based flood mapping. Different data sources also require other suitable methods for research. Data archives covering the Netherlands during the flood in 2021 will influence the choice of data. The case study for this thesis will use the Sentinel-1 stack that is available in both VV and VH polarization. Additionally, 3 acquisitions from Capella Space are available during the peak of the flood. Using this case study and the corresponding datasets, different methods and parameters are tested to find an optimal approach for estimating the Water-Land boundary. Intermediate results and user needs will define the methods that are pursued for the final results. Water level gauges, an hydrodynamic model with a time component, aerial imagery and LiDAR-derived water lines are taken as validation data.

1.4. Outline

The research of this thesis will include various components described in the remaining chapters. Chapter 2 will provide general background knowledge about Synthethic Aperature Radar. The chapter highlights the important parts of the technique and how it can be used for flood monitoring. The end of the chapter will describe the considered data products in further detail. Chapter 3 describes the methodology for the research. The chapter explains the choices for considered and pursued methods with the underlying theories. Finally, the validation of the results is described. Chapter 4 is dedicated to showing the results of the proposed methods and Chapter 5 entails the discussion of the results. Strengths and shortcomings for all methods are discussed, together with an evaluation of the accuracy. Chapter 6 summarizes the important conclusions that can be extracted from the discussion and provides recommendations for future research.

2

Principles of Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a type of RAdio Detection And Ranging (RADAR), an active remote sensing (RS) technique. The most common techniques in remote sensing are passive optical imagery, examples are images acquired by Landsat from the U.S. Geological Survey (USGS) or Sentinel-2 from the European Space Agency (ESA). Their range of applications goes from mapping land and sea surface temperatures to estimating properties of vegetation or atmosphere like clouds and aerosols. Passive remote sensing uses the Sun's energy to sense the electromagnetic radiation (EMR) emitted by the Earth that reaches the sensor of the satellite. As mentioned, optical sensors are mostly passive techniques. These sensors operate in the Near Infrared (NIR), Infrared (IR), Thermal Infrared (TIR) and Visible (VIS) part of the EMR. On the contrary, active remote sensing techniques provide their own energy source making them independent of the Sun's energy. An electromagnetic wave is send out to the Earth's surface which will reflect part of the signal back to the sensor. Some applications for active remote sensing techniques are: measuring the topography of glaciers, forest canopies or sea surface structure. Both techniques are displayed in Figure 2.1. SAR is an active remote sensing technique characterized by its coherent, microwave and weather independent way of imaging with high spatial resolution. Microwaves operate in a different part of the EMR, the wavelength is larger therefore penetration through the clouds is possible. This makes the technique not only independent of the Sun's energy but also weather independent. This sensing technique differentiates SAR-systems from other remote sensing techniques and makes it interesting for further research into the wide range of applications. In this chapter the principles of SAR will be discussed, the applications to floods, and the SAR data sources that will be used for this thesis [14].

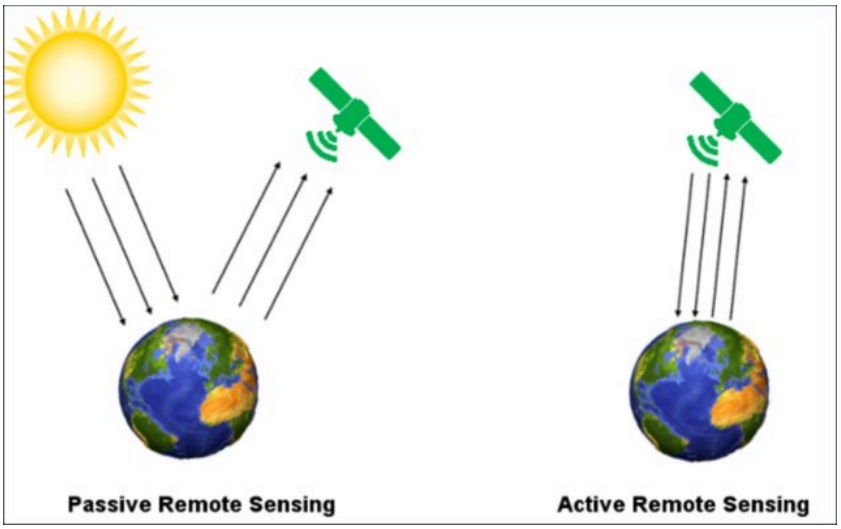


Figure 2.1: The difference between passive and active remote sensing sensors [15].

2.1. The basic concepts of SAR

As described above, SAR is an active remote sensing technique that operates in the microwave part of the electromagnetic spectrum. Active remote sensing techniques send out electromagnetic pulses that will interact with the Earth's surface and record the energy that is reflected back to the sensor, referred to as backscatter. The distance to the surface can be computed by recording the amplitude of the reflected signal and the time between the transmitted and received signal. A radar is usually affixed to a moving platform like an airplane or a satellite. A signal is emitted to the surface, which will scatter a part of the energy back to the satellite. As the radar is its own illumination source, SAR satellites are independent of natural processes like day or night [16].

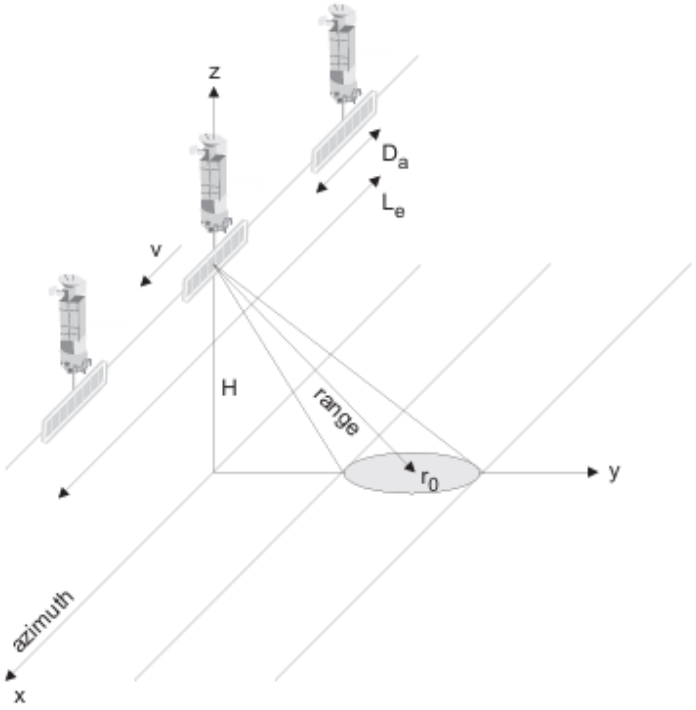


Figure 2.2: The concept of Synthetic Aperature Radar.

The radar moves in the flight direction which is the azimuth or along-track direction and it images a swath along the ground. The antenna points towards the ground with a certain incidence angle and perpendicular to the flight direction, referred to as the range or across-track dimension. The acquisition geometry seen in Figure 2.2 is standard for SAR imaging. The range resolution, δ_r, is inversely proportional to the signal bandwidth (B) and given by

$$\delta_r = \frac{c}{2B}, \tag{2.1}$$

where c is the propagation speed of electromagnetic waves. The range resolution is the minimum range separation (in meters) for points to be distinguished by the radar in the range dimension [16].

Resolving targets in the range dimension is done similarly in almost all radar systems. The SAR technique is distinguished from other radar systems as it resolves targets in the azimuth dimension. In general, the antenna beam has a spread in azimuth related to the wavelength (λ) and the antenna length in azimuth direction (D_a). The azimuth resolution will therefore depend on the spread and the range distance between the antenna and the scatterer (r_0):

$$\delta_a = r_0 \frac{\lambda}{D_a}. \tag{2.2}$$

From Equation 2.2, can be deduced that the azimuth resolution will be higher (thus a smaller value) with a longer antenna. When looking at the wavelength of a C-band radar, which is 5 cm, an antenna of over 4 km is needed to obtain a spatial resolution of 10 m. In terms of practicality, this is unfeasible in space, hence the introduction of the synthetic aperture concept. Using the motion of the satellite, a target will be visible from different positions as the satellite moves in the flight direction, simulating one large antenna as shown in Figure 2.2. The synthetic antenna combines the sequence of acquisitions to simulate the larger antenna, thus obtaining higher azimuth resolution:

$$\delta_a = r_0 \frac{\lambda}{2L_e}, \tag{2.3}$$

where L_e is the maximum length of the synthetic aperture for a target at slant range distance r_0. L_e is limited by the illumination time of the target by the radar beam as follows:

$$L_e < \frac{\lambda R}{D_a}. \tag{2.4}$$

As a result, the maximum azimuth resolution for a synthetic aperture depends on the antenna dimension D_a. Deducting from Equation 2.5, the azimuth resolution is not dependent on the target's range position nor the wavelength for a SAR satellite system [16] :

$$\delta_a > \frac{D_a}{2}. \tag{2.5}$$

The antenna of the satellite measures three things from the received pulses: intensity *I*, phase ψ and the time it takes from sending the signal to receiving the signal. The information within a single pixel is the sum of all the reflections from the corresponding surface area. The intensity of the reflections is represented by the amplitude: $A = \sqrt{I}$. The images are in a line-pixel coordinate system, where each pixel *P* contains the amplitude and phase in a complex-valued number

$$P = Ae^{j\psi}. \tag{2.6}$$

The amplitude and phase have different units. To derive both by the same formula, they are expressed by a complex equation. The complex equation consists of a real and imaginary part that represents the two different components. This is a conventional way of mathematically combining two units into one complex signal. The real part of the equation is denoted as $Re\{P\} = A\cos(\psi)$ and the imaginary part as $Im\{P\} = A\sin(\psi)$. The complex plane displaying the relation is shown in Figure 2.3, where the coordinates of a point consists of a real and imaginary part. The real and imaginary parts are related to the amplitude and phase, respectively [17]

$$A = \sqrt{(Re\{P\})^2 + (Im\{P\})^2}, \tag{2.7}$$

$$\psi = \arctan\left(\frac{Im\{P\}}{Re\{P\}}\right). \tag{2.8}$$

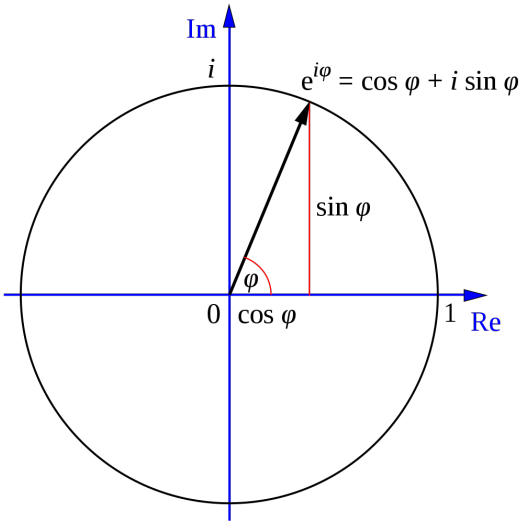


Figure 2.3: Relation between the real and imaginary parts of the signal presented as a point in the complex plane.

Above a C-band radar was mentioned. Different wavelengths used for SAR systems are known as bands. The bands are assigned letters, the most common ones are X, C, L and P. The letter designation was invented because of their war applications in the early days. In this way, the opponent would not know what people were talking about. Table 2.1 shows the different bands and their respective wavelengths and frequencies. The wavelength determines the interaction of the signal with the surface and the penetration depth into a medium, therefore an important parameter that is used differently depending on the application. For example, X-band operates at a wavelength of 3 cm which will interact mostly with the leaves of a tree, while L-band operates at 23 cm which will interact mostly with the trunk of the tree [14].

Table 2.1: Different SAR-bands and their corresponding frequency and wavelength.

Band	Frequency	Wavelength
X	8-12 GHz	3.8-2.4 cm
C	4-8 GHz	7.5-3.8 cm
L	2-4 GHz	30-15 cm
P	1-2 GHz	100-30 cm

2.2. Polarimetry and Scattering Mechanisms

Radar is able to receive signals in different polarizations. Polarization is the term for the orientation of the plane in which the transmitted electromagnetic wave oscillates, see Figure 2.4. The orientation of the plane can take any angle, however SAR-systems tend to transmit linearly polarized waves. A distinction is made between Horizontal polarization (H) and Vertical polarization (V). The radar can collect multiple polarizations by controlling the polarization in the transmission and receiving paths. In this way, signals emitted in the vertical plane can be received in the horizontal plane and vice versa. These signals are indicated by VH/HV polarization, or generally referred to as cross-polarization. The same nomenclature is used for single-polarization signals that transmit and receive in the same plane, this is called VV or HH polarization [18].

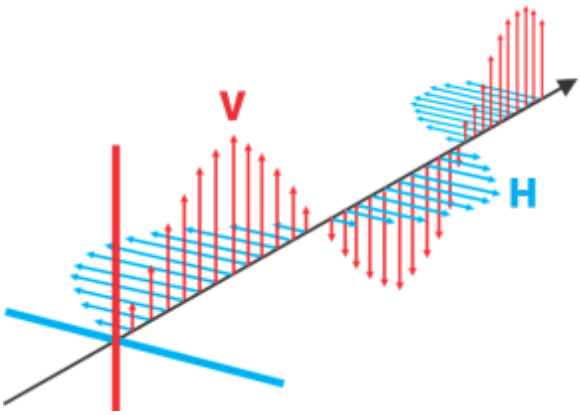


Figure 2.4: Planes of Vertical and Horizontal polarization.

Dual-pol instruments are satellites that transmit either H or V polarization and receive H and V simultaneously. Quad-pol instruments can transmit H and V while alternating the pulses, and receive both H and V simultaneously. How much of the polarized signal is received depends on the physical properties and geometric structure of the surface, and how these interact with the microwaves that have been transmitted from the satellite. When a surface target is composed of varying orientation such as leaves, branches or flowers, the waves will be scattered by these elements creating waves with varying phase and polarization [14].

Radar polarimetry is the study of multiple polarimetric returns to infer information about the surface. There are two approaches to studying this concept. The first is to use theoretical models that predict how polarized signals interact with different media. The second are observations made with remote sensing instruments that can reveal the signatures of a certain land cover. By making a color composite of a scene with different polarizations one can infer the surface characteristic of that area, this is displayed in Figure 2.5.

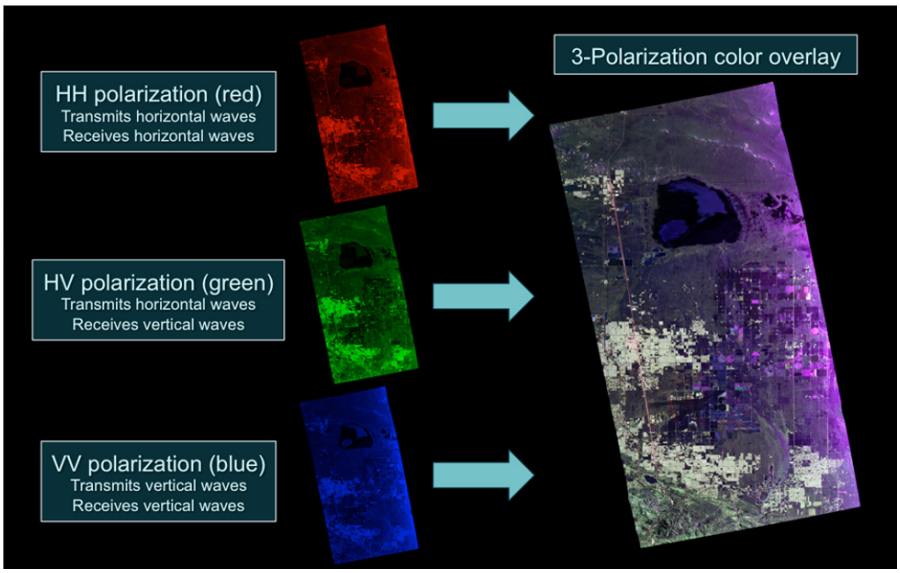


Figure 2.5: The combination of different polarizations in an RBG image to extract surface characteristics [19].

Based on the polarization and the strength of the signal, information of the surface can be examined as this is related to the interaction of the signal with the ground. This interaction is also referred to as scattering. Figure 2.6 shows the three main types of scattering: Rough Surface, Volume and Double Bounce Scattering [14].

- Rough Surface Scattering: related to bare soil or water, more structural surfaces will lead to increased scattering, strong VV scattering.
- Volume Scattering: related to leaves and branches of a canopy, strong VH or HV scattering.
- Double Bounce Scattering: related to buildings, tree trunks or inundated vegetation, strong HH polarized signal.

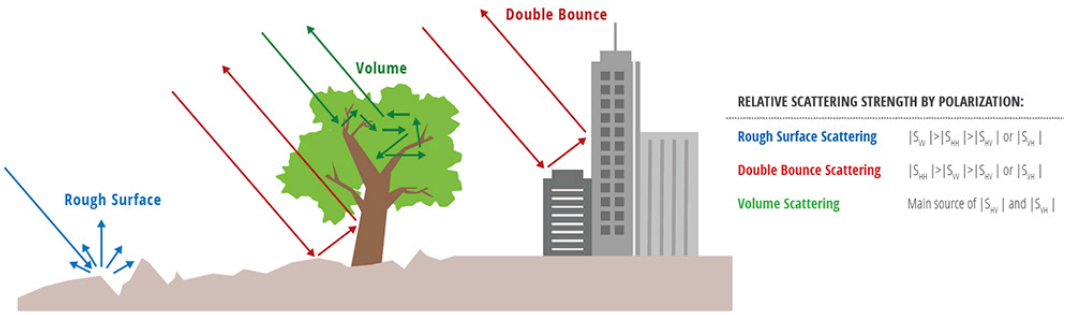


Figure 2.6: Scattering Mechanisms [14].

Concluding, different surface characteristics will have different effects on the reflected signal. Depending on the wavelength, the surface will affect the signal in such a way that scattering occurs. The interaction between the surface and the signal with natural or man-made objects will affect the penetration depth of the signal.

2.3. InSAR

Interferometry is the way information is exploited from the interference of waves. Satellite radar interferometry (InSAR) is a technique that uses multiple SAR acquisitions to mathematically correlate them to obtain surface information. Single acquisitions return a single wave for each pixel with an amplitude and an absolute phase. The amplitude is related to the strength of the radar signal, whereas the phase is a fraction of a completed sine-wave cycle (or a single SAR wavelength). The phase is inherent, so temporal changes can be detected by computing the phase difference between two consecutive acquisitions. The signal has to travel to the surface scatterer and back to the sensor (two-way travel). When the location of the scatterer changes, the distance changes, hence the phase changes between the transmitted and received signal. When the phase difference between two waves or two acquisitions is a fixed phase shift over time, this is considered to be coherent. Coherent waves imply that the frequency and wavelength must be the same. The phase differences are the basis for the interferometric phase and InSAR. A complex interferogram can be created by complex conjugated multiplication of two aligned (referred to as coregistred) images, computed by

$$P^{1,2} = P^1 P^{2*} = A^1 A^2 \exp\left(i(\psi^1 - \psi^2)\right), \tag{2.9}$$

where P^{2*} is the complex conjugate of the second image [17]. The coherence can be interpreted as a measure of similarity between two images that can form an interferogram. It compares the complex values of both the amplitude and phase of the two images and combines them into a normalized value ranging from 0 to 1. Where, 0 indicates low coherence and no similarities between two radar images, and 1 indicates a high coherence and perfect similarity between the images.

Pixels in SAR images are composed of the scatterers that are inside. Due to stochastic scattering processes, it is expected that each pixel shows a small decorrelation. This is even expected for perfect scenarios when no changes occurred and the SAR acquisition was at the same angle and location [18].

2.4. Speckle

Aspects that make SAR images generally less favorable compared to optical images are the resolution and noise levels. Noise in SAR images is a granular noise, portrayed as a salt and pepper pattern, and referred to as speckle noise. Speckle is a phenomenon that appears due to the constructive and

destructive interference of reflected waves by many scatters on the surface [16]. Inside each resolution cell, there are several scatterers that contribute to the final intensity. These scatterers (or reflections) within a cell add up together as a random sum, some enhancing and some destroying each other. This creates positive or negative interference. The summation is unique for each cell, which can cause two neighboring pixels with similar surface characteristics to have a different intensity value. An example of constructive and destructive interference is portrayed in Figure 2.7.

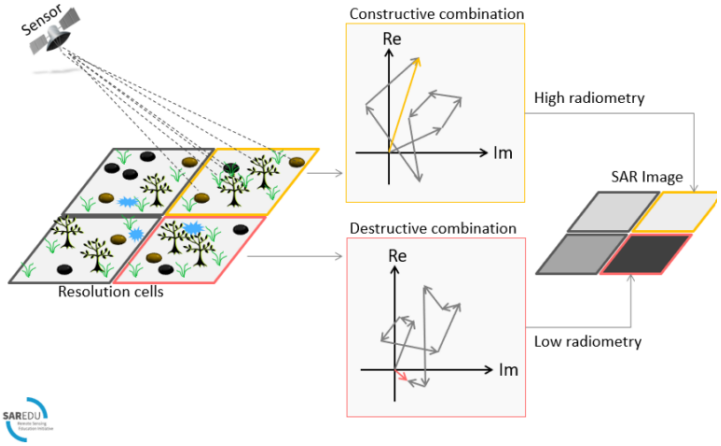


Figure 2.7: Generation of speckle due to constructive and destructive interference [18].

The amount of speckle visible in an image will reduce the capability of object distinction and is therefore an important challenge during processing. To increase the ability of SAR for flood mapping, a despeckling filter can be applied to decrease the amount of speckle in a SAR image. Input data is significant for the performance of many analysis methods, therefore despeckling of SAR images can be crucial to finding the most accurate flood extent delineation. Traditional despeckling approaches are either in the spatial or temporal domain, although recently published Deep Learning techniques try to optimize both while limiting the resolution loss. Three different, yet widely used, despeckling techniques are described below.

2.4.1. Temporal Averaging

Temporal averaging is one of the traditional despeckling processes. This technique will exploit the temporal correlation between the images, and thereby enhancing the Signal-to-Noise ratio to improve image quality. The pixel values from different image dates are combined and averaged, this will reduce the random noise in the pixel and preserve the coherent signal. As a consequence, temporal resolution will decrease. Applications for this technique are for a static Area Of Interest (AOI), where different dates obtain the data of interest.

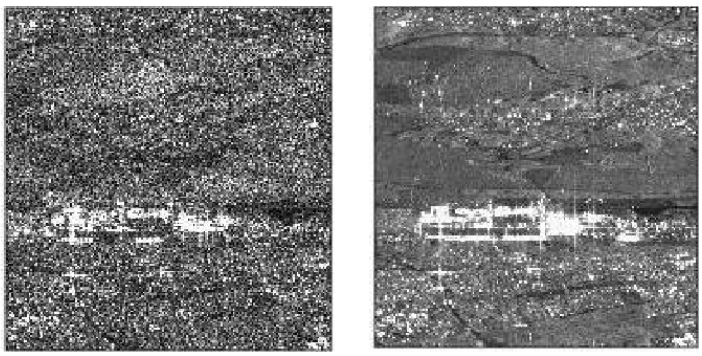


Figure 2.8: Example of the difference between the noisy Sentinel-1 SAR data and the arithmetic mean of the image stack [20].

2.4.2. Refined Lee

The Refined Lee filter is the enhanced Lee filter and is one of the most widely used despeckling filters. The Lee filter is based on local statistics calculated using a square window. The center pixel will then adapt to the values of all neighboring pixels. The Refined Lee Filter is based on the Lee filter, enhancements of the method are focused on spatial despeckling by estimating local statistics, among others the mean and variance. Using a sliding window for each pixel, a weighted average is computed for smoothing of the image by development of a K-Nearest Neighbor (KNN) algorithm to adjust the number of neighboring pixels. The result will contain a loss of spatial resolution. It can be seen in the image by the loss of finer details such as sharp edges or textures within the image [21].

(a) (b) (c)

Figure 2.9: Example of the implementation of the Refined-Lee filter, from left to right respectively: original SAR image, Refined Lee filter with 5x5 window and Refined Lee filter with 7x7 window [22].

2.4.3. Deep Learning-Based Despeckling

Lastest despeckling techniques are based on Deep Learning models. Deep Learning is a subset of Machine Learning, which is a subset of Artificial Intelligence. It utilizes artificial neural networks to process data tasks, inspired by the human brain. A few examples of widely used neural networks for despeckling are: Convolutional Neural Networks, Autoencoders and U-Net [23]. Neural networks can be trained by supervised (combination of speckled and despeckled data) or self-supervised learning (set of speckled image pairs or a single image) techniques. The trained networks learn to filter out the speckle noise and recreate a clean, speckle free, image.
MERLIN (coMplex sElf-supeRvised despeckLINg) is an example of a self-supervised training strategy that is obtained with Deep Learning. It is based on the separation of real and imaginary parts of the single-look complex SAR images, offering a straightforward way to train different kinds of deep despeckling networks. The networks take spatial correlations into account with an SAR transfer function that is specific for each sensor and imaging mode. It requires only 1 image, but the method can exploit large archives, making it possible for large-scale training of despeckling networks. The applicability of a pre-trained network requires different fine-tuning parameters for each type of satellite. When similar fine-tuning parameters are used, the performance of the model is not consistent for each satellite. The despeckling method is based on pre-defined training datasets that predicts the amplitude per pixel. This results in limited resolution loss in both temporal and spatial domain [24].

Despeckling reduces the Signal-to-Noise ratio, thereby improving visual interpretation of images and identification of structures and edges. One of the disadvantages of despeckling is a shrink in distribution, resulting from the reduction of noise. Consequently, this shift in distribution contributes to the smoothing of the data, making it more challenging to distinguish surfaces with similar backscatter profiles.

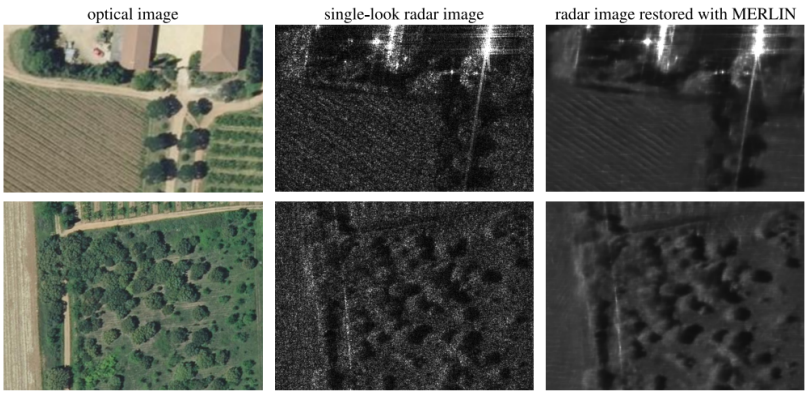


Figure 2.10: Example of MERLIN Despeckling [24].

2.5. SAR-based Flood Mapping

SAR-based flood mapping in different areas requires other approaches. The signal of the satellite reflects differently on each type of surface. For this thesis, an optimal methodology is established for the area of the Netherlands that has been affected by the 2021 flooding event. The main focus will be on rural areas with vegetated soil as the event occurred during a wet summer season in the Netherlands, hence high vegetation in rural areas.

During the flood event in 2021, both cities and rural areas were affected. Different characteristics of the Earth's surface can be distinguished by the amplitude and phase of the signal. The physical and electrical properties of the target will influence the interaction with the signal and how much is reflected back to the sensor. Parameters such as geometry, roughness and dielectrical constant of the surface are examples of such properties.

2.5.1. Amplitude

The amplitude or backscatter information is the most widely used part of the SAR signal for flood mapping. This is because the SAR signal displays specular reflection over water bodies. Specular reflection scatters the SAR signal away from the sensor, resulting in a low backscatter value. This interaction is shown in Figure 2.11. As can be seen in the figure, contrasts are large between the water and the dry conditioned surface. This contrast can be used to extract the flood extent at the time of acquisition. Specular reflection can be influenced by parameters such as wind, precipitation, and emergent vegetation, complicating water detection [25].

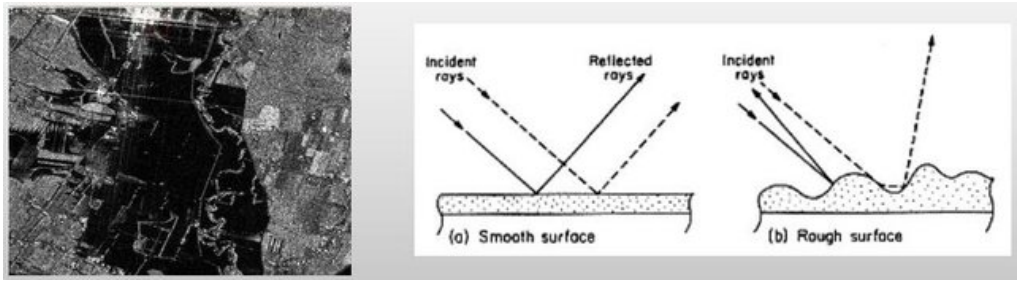


Figure 2.11: SAR intensity band image showing low intensity over water bodies. The diagram displays the water-signal interaction based on the surface roughness [26].

In this thesis, multiple parts of the radar signal can be used to analyze the information that can be contributed by each component. Amplitude data is the most frequently used in papers, because the specular reflection phenomenon is the most prominent change during a flooding event. However, polarimetric data as well as InSAR coherence are considered to be important parameters that have been shown to provide complementary information about the flooded area [27].

2.5.2. Coherence

Recent studies have shown much progress in the near-real time flood mapping approaches. However, detection of flood inundated areas remains a complex procedure due to vegetated and urban areas having ambiguous radar signatures. Detection of water in urban environment is more likely to be detected by using the high sensitivity of interferometric coherence to small changes in the area.

In urban areas, double bounce effects are predominantly present as they appear due to the presence of buildings. This is explained in section 2.2. Detecting surface water rise in front of buildings during a flood event, can be done by looking for an increase of the double bounce effect. These effects are mainly seen in the co-polarization channel. Studies have shown that this effect is visible when the angle between the flight direction and the street alignment is small, the effect decreases for larger angles. Consequently, the increases in double bounces due to floodwater are not significantly present in the images when the buildings are not built parallel to the satellite flight direction [28]. When dealing with complex urban environments, flood water detection is not fully optimized while using only the backscatter intensity. To further look into these effects, InSAR coherence data can be integrated. Coherence is the normalized cross-correlation between two consequent images, therefore related to the changes in spatial arrangement of the scatterers in a pixel of the image. This also includes geometric changes. Coherence in urban areas can provide different information compared to backscatter and thus further help to extract the presence of flood water. Coherence is expected to be high throughout the entire year, hence when floodwater is present it can significantly decrease the coherence [27]. The coherence can also be used for land cover information and to reduce false alarms that are caused by volume scattering on vegetated soils.

2.5.3. Polarization

In addition to coherence and backscatter, changes in polarization can also provide complementary information in vegetated areas. Cross-polarization is more sensitive to volume scattering compared to co-polarization, which is more sensitive towards double-bounce scattering. Studies from Pelich et al. have shown that combining the information from both polarizations increased the accuracy of the flood maps from 75.2% to 82.9% [27]. Contributions and differences of both polarizations for a coherence map are displayed in Figure 2.12, a flooding event in Houston. The image displays the change in coherence during a flooding event for both VV and VH polarizations. As can be seen in the image, different areas that can be detected by the use of the different polarization types, thereby making the flood extent delineation more complete when both polarizations are used.

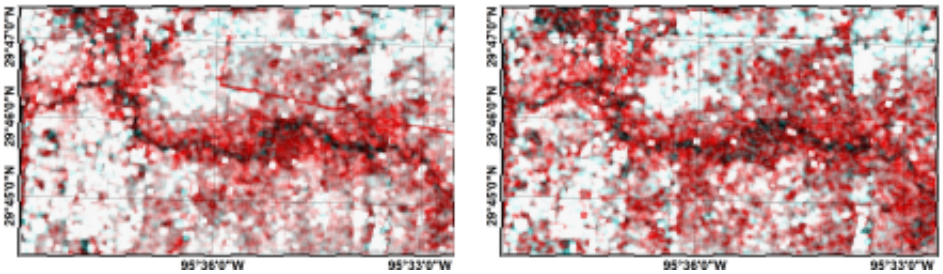


Figure 2.12: VV and VH InSAR Coherence RGB composites of pre- and co-event image pairs during a flooding event in Houston, USA. The images display the difference in InSAR coherence with: R=pre-event coherence image pair and G=B=co-event image pair. [27]

The usage of different components of the SAR signal depends on the studied area. In ”standard” flat and bare surface areas, floods generally lead to a significant decrease in backscattering. The amount of decrease is related to the change in surface roughness from bare soils to water. The actual roughness of the water is related to physical parameters (wind or debris) and system parameters (polarization, operating frequency and incidence angle). Concerning standard surface conditions, several studies have shown that co-polarization generates the most accurate flood extent. The highest contrast between water and land is achieved using HH polarization, followed by VV polarization [29].

When the affected area includes the presence of vegetation or buildings, the backscattering becomes more complex. Flooded vegetated areas include scattering of the surface, volume scattering from vegetation and double bounce scattering from the trunks. The method should be well-defined in order to limit underestimations of the flood extent. In general, cross-polarization channels are less affected by volume scattering and larger wavelengths (L-Band) can penetrate deeper into vegetation. Methods using additional data from cross-polarization channels and larger wavelengths have shown to yield more accurate flood extents in the presence of vegetation [29] [27].
If the area includes man-made structures such as buildings, double bounce scattering is dominant in normal conditions. Urban areas are coherent areas, therefore usage of coherence can indicate the presence of flood water. This principle relies on the expectation that the double bounce scattering intensifies when flood water is present. As explained earlier, the behavior of the signal is related to the orientation of the buildings with respect to the flight direction. Despite the complexities, several studies have shown that the use of coherence data improves the accuracy of flood extent mapping in urban areas [30].

2.5.4. Methods

There are various methodologies when it comes to delineating flood extent. The chosen approach should depend on the area, requirements of spatial and temporal resolution and availability of the SAR data. The most common approaches are: Tresholding, Change Detection and Time Series Analysis.

Thresholding methods are widely used in this field, the differences in backscatter between water and dry land are often sufficient to find a suitable distinction. Histogram thresholding is a fast and rapid technique for flood mapping. The selection of the threshold is a critical step that will define the results of the outcome.

Change Detection approaches are a little more complex, the method focuses on the comparison between at least two images: at least one pre-event and co-event image. Choosing the correct pre-event image can be difficult because of the dynamic changes of backscatter over time [31].

Other approaches focus on using the image archives and try to model the behaviour of objects in SAR images over time. Time Series Analysis (TSA) looks further into the normal situation of the unflooded area and tries to estimate the trend. This approach requires proper temporal modeling of the backscatter. In order, the approach is increasingly complex yet provides more accurate results compared to Thresholding and Change Detection. Different methods considered for this thesis are described in the next section.

2.6. Area of Interest

For this thesis, the selected AOI is located in Limburg, North of Maastricht. The AOI has been significantly affected by the floods in 2021, inhabitants of both Bunde and Meerssen had to evacuate the area because of the high water levels. The villages are located close the Geul stream. Generally, the stream has an outflow into the Meuse river and flows further into the country. During the event, the stream could not flow into the river as a consequence of high water levels in the Meuse. As a result, the stream water reached outside the banks and into the villages. The map in Figure 2.13 displays the chosen AOI of the case study, the maximum flood extent computed by the ENW report [8] and the permanent water bodies of the Geul stream and Meuse river.

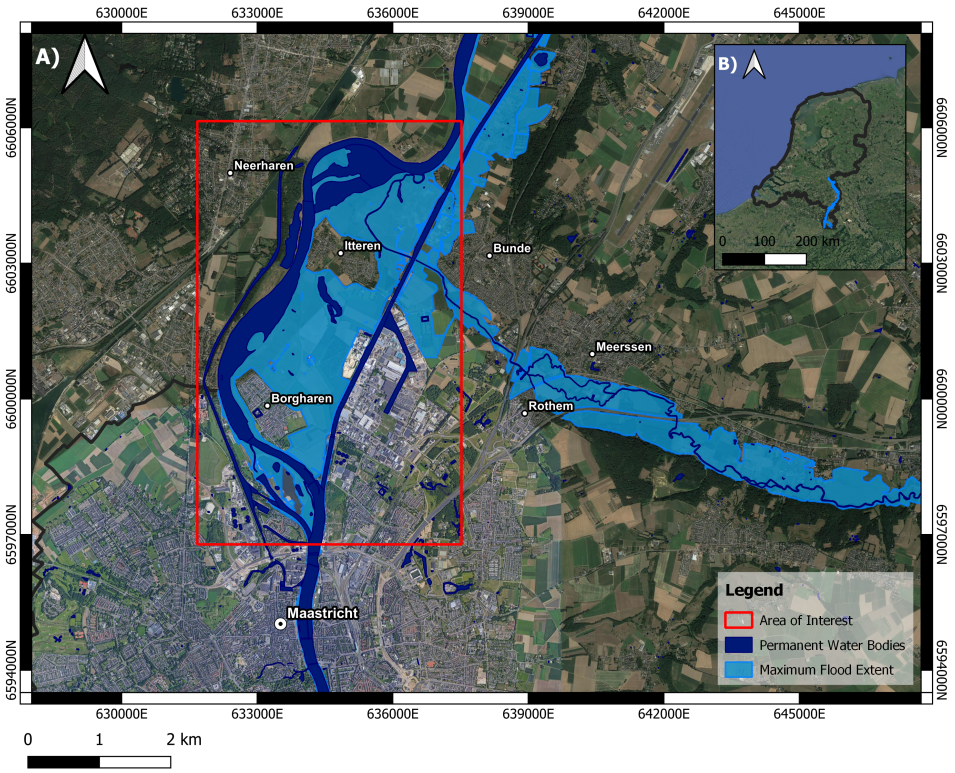


Figure 2.13: The Area of Interest and the maximum flood extent of the areaare shown on the map, where A) is on a local level and B) indicates the total affected area of the flood on a national level. In A) the Meuse river and the Geul stream are highlighted to display the locations of permanent water bodies in the area.

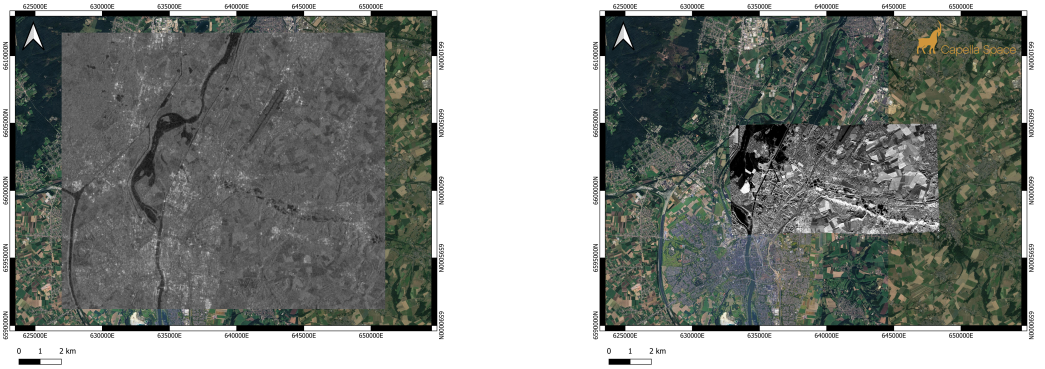
2.7. Data products

2.7.1. SAR products

This thesis focuses on using different radar sources and practical flood detection techniques. In the recent years, new innovations have caused a rise of very high resolution (VHR) or sub-meter resolution commercial SAR imagery products. These satellites fly in a multiple satellite-constellation at an altitude of only a few hundred meters. The aim is to capture the Earth in a range of hours which can be ordered on-demand. The launch of these satellites has started since 2017 with companies such as Iceye, Capella Space and Umbra Space. VHR data can obtain resolutions up to centimeter level. With their higher revisit frequency, spatial resolution, and acquisition flexibility, these data products will be able to complement the open source data. For our case study, Capella Space will be used as they have archived images from the flood event period over the area of Meerssen, Limburg. Capella Space provides images with an X-Band radar and spatial resolution up to 50 cm [12]. Other radar sources contain archives starting from 2015 and have been acquiring data consistently over the years. Therefore, the second data choice is considered to be of medium resolution with 5 m x 20 m pixels and is an open data source: ESA's C-Band SAR Sentinel-1 with data starting from 2015. With long historical data archives, this data type is favorable for Time Series Analysis [32]. Both data sources are shown in Figure 2.14, showing the image acquired during the flood event.

Table 2.2: Data products used in this thesis.

SAR System	Band	Frequency & Wavelength	Polarization	Acquisition Date + Time
ESA Sentinel-1	C-Band	5.405 GHz & 5.55 cm	VV + VH	15/07/2021 05.50 UTC 17/07/2021 17.30 UTC 18/07/2021 17.25 UTC
Capella Space	X-Band	9.65 GHz & 3.1 cm	HH	15/07/2021 23.30 UTC 17/07/2021 20.15 UTC 18/07/2021 20.20 UTC



(a) Sentinel-1 image.

(b) Capella Space image. © 2023 Capella Space. All Rights Reserved [12].

Figure 2.14: SAR images from both data sources used in this thesis acquired in July 2021 over Limburg, the Netherlands.

Capella Space

Capella Space can provide sub-meter VHR (0.7 m) X-Band SAR data, in a multi-satellite constellation. The choice for on-demand data can be preferred over the longer revisit time of 6-12 days from Sentinel-1. A revisit time of 6 days may not catch the peak of the flood because of the short duration of floods in the Netherlands. Higher temporal resolutions cannot be found with current public providers but commercial providers are further ahead. The Capella Space SAR constellation is an on-demand system that is able to collect data in the order of hours for emergencies during and after floods. The company is based in the USA and are committed to delivery of high quality and high-resolution SAR imagery with unparalleled frequency and timing of images [12].

Currently there are six operating satellites in a low Earth orbit (altitude of less than 1000 km) [12] [33]:

- Capella-9-WHITNEY with an inclination of 44^o
- Capella-10-WHITNEY with an inclination of 44^o
- Capella-11 (ACADIA-1) with an inclination of 53^o
- Capella-13 (ACADIA-3) with an inclination of 53^o
- Capella-14 (ACADIA-4) with an inclination of 45.4^o
- Capella-15 (ACADIA-5) with an inclination of 97.0^o

Several data products from Capella Space are available: Single Look Images (SLC), Geocoded terrain corrected (GEO) and Geocoded Ellipsoid Corrected (GEC) images. SLC images preserve both amplitude and phase information, GEO images are multi-looked and geocoded, and GEC images are multi-looked, geocoded and projected onto the WGS84 ellipsoid.
As the GEO data product optimizes the geolocation accuracy, this product is used in this thesis. Preprocessing of the GEO-data contains despeckling using the SNAP toolbox with a 3x3 window Lee filter and converting the intensity values to dB.

Sentinel-1

Flood events in the Netherlands have a time span ranging from a few hours to a few days. The C-band radar carried by Sentinel-1 has a revisit time of 6 days and has showed promising results for flood mapping in previous studies [29]. Since we are looking at an event that lasts for a short time period, a short revisit time is preferable. Higher spatial resolution satellites generally have longer revisit times, such as TerraSAR-X or RADARSAT-2 have a revisit time of 11 days and 24 days respectively. At the current time, Sentinel-1B is no longer in space due to the anomaly detected at the end of 2021 [34]. However, for this case study both Sentinel-1A and 1B were still recording at the time. For future research the new satellite Sentinel-1C has been be launched in December 2024, preserving the temporal resolution as proposed above. Other capabilities of Sentinel-1 are the medium spatial resolution of 5 m x 20 m that is acquired with a very wide swath width of 250 km. Additionally, Sentinel-1 has a systematic acquisition of dual-pol SAR data with VV and VH polarization available over the Netherlands. The images used in this thesis are Single Look Images (SLC) and Ground Range Detected (GRD) images. SLC images preserve both amplitude and phase information whereas GRD images only contain amplitude information. The advantages of GRD images lie in quick usage and accessibility from sources such as the Google Earth Engine.

2.7.2. Digital Elevation Model (DEM)

In order to extract water levels from the flood extent, the terrain heights are used. For this research, the study area is located on the border of the Netherlands and Belgium. The local DEM of the Netherlands and the local DEM of Belgium are therefore used.
In the Netherlands, the local DEM is the Actueel Hoogtebestand Nederland (AHN). It contains the elevation data covering the entire country with a vertical accuracy of 5 cm. The elevation data is LiDAR dataset, collected from helicopters and aircraft. The data is taken between 2020 and 2022, this includes the period of the flood event [35]. The dataset used for this thesis is the AHN4, containing 0.5 m Digital Surface Model (DSM) and 5 m Digital Terrain Model (DTM). The DTM represents the elevation of the Earth's surface and the DSM represents the elevation of Earth's surface including objects such as buildings or trees [36]. The measured heights are referenced to Normaal Amsterdams Peil (NAP).
In Belgium, the local DEM used is the 1 m Digitaal Hoogtemodel Vlaanderen II (DHMV2) [37]. The elevation data is also acquired by LiDAR data and covers the Flemish region of Belgium. The data is taken between 2013 and 2015. The dataset contains both a DSM and DTM with a vertical accuracy of 5 cm. The measured heights are referenced to the Tweede Algemene Waterpassing (TAW). The reference points of both DEMs are not aligned, therefore a correction must be made to use both. The TAW reference point is known to lie approximately 2.33 m beneath the NAP reference point. A correction of 2.33 m is therefore made to the TAW dataset to align the DEMs [38].

2.8. Validation Sources

2.8.1. Water Level Measurements - Landelijk Meetnet Water

Rijkswaterstaat is part of the Dutch Ministry of Infrastructure and Water Management, is responsible for measuring water levels using gauges in the rivers. There are more than 450 locations along the coast and are spread across the biggest rivers and streams of the Netherlands [39]. On these locations, research is performed on the water quality, biology, chemistry and dynamics of waters. This information is crucial when it comes to the safety of the country during flood events. The department of Landelijk Meetnet Water (LMW) is responsible for the water level gauges to continue measuring accurately. The gauges measure the water levels at specific locations every 10 minutes, with historical data dating back to 2020. Within this archive, the measurements during the flood can be collected. There are four locations that the water level estimations from the SAR-based flood exents can be validated to. The gauges that are intersecting the AOI are selected and shown in Figure 2.17a. The coordinates of the gauges are often located on the sides of the river whereas the measurements represent the water level at the center line of the river. In order to properly validate the data sources, the locations are projected to the center line of the river. The assumption can be made that within a radius of 50 meters from the gauge, the water levels are be constant. Thus, on the intersecting locations of the SAR-based flood extent and within the 50 meters radius, the water level estimates can be compared to the the measured value at the center line of the river. Figure 2.15 shows the LMW gauge at Bunde.

Figure 2.15: Example of the LMW gauge near Bunde, projected onto the center line of the river and the surrounding 50 meter radius that is assumed to have a constant water level.

2.8.2. Hydrodynamic model - Deltares

Specialists from the knowledge institute of Deltares have computed an hydrodynamic model that simulates the developments from the flood event of the Meuse river over time. The model computes the region of the flood on 20 m x 40 m grid cells that contains information about several variables. Among others, the water level and water depth. The simulated water depth of the model is displayed in Figure 2.17b. As the model contains a temporal component, it is possible to compare data to both the results from the Sentinel-1 acquisition and the Capella Space acquisition. The validation will be performed to evaluate the results of the computed flood extent from the SAR satellites to the modeled results and indicate whether the timing of the model corresponds to the timing of the satellite acquisitions. If the temporal component does not align well with the SAR-based flood extent, this could indicate that the inflow or outflow is not yet simulated accurately in the model. The latter comparison is of importance to exploit the possibilities of calibrating the models with SAR data.

2.8.3. Aerial photographs from Waterschapshuis

Under instructions of governmental institutions, the flood event in Limburg has been captured by airplanes with aerial images and LiDAR measurements.The airplanes flew for four days along the peak of the flood to capture the maximum water levels in the riverbeds. The first flights were executed from Eijsden to Susteren on the 16^{th} of July, then Susteren to Beesel on the 17^{th} of July, Beesel to Mook 18^{th} of July and the last flight was Mook to Hank on the 19^{th} of July. Figure 2.16a shows the entire area that has been captured in the images [9]. The data has been published by Het Waterschapshuis and the Waterschap Limburg by means of a Storymap, displaying the difference between spring of 2021 and the flood event (see Figure 2.16b).

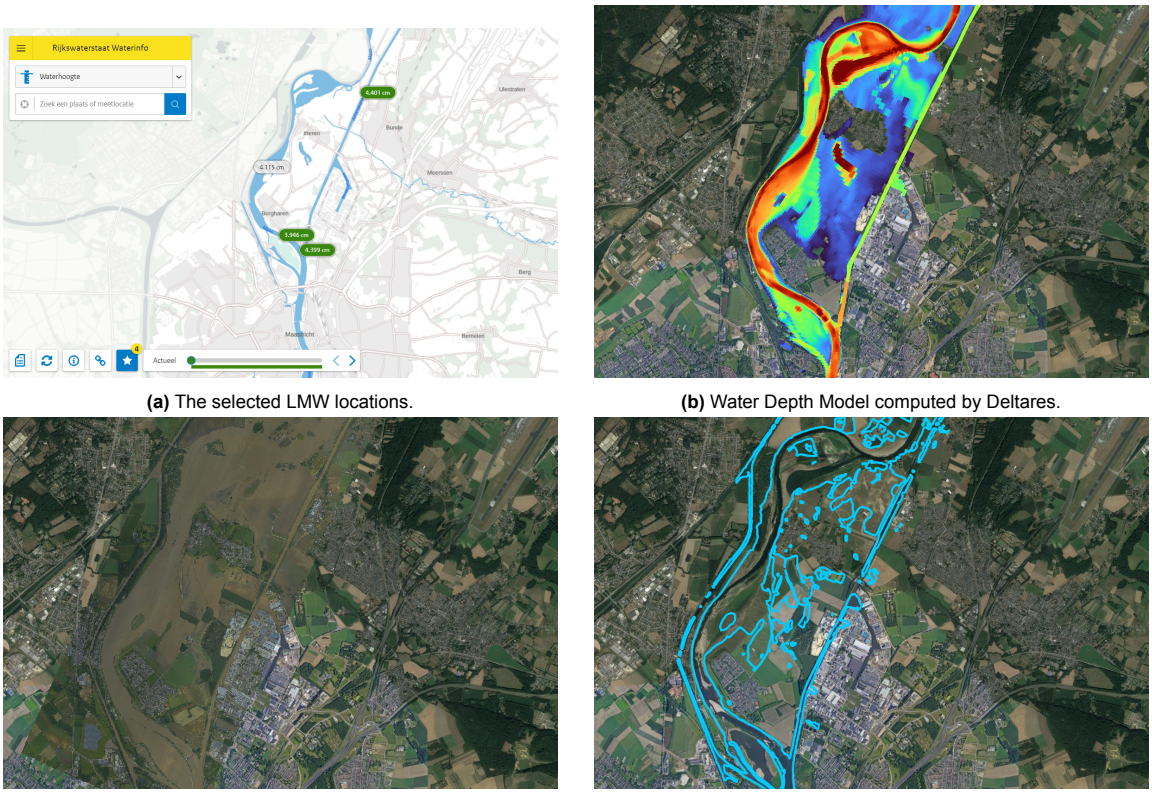
(a) The full area of the Meuse river that has been captured.

(b) Differences between spring and during the flooding in July 2021.

Figure 2.16: Aerial images taken by airplanes during the flood event, displayed on the published Storymaps [9].

2.8.4. Water-Land Boundary lines

LiDAR measurements have also been obtained by the Waterschapshuis. From the LiDAR data and additional data provided by civilians, the flooded area has been mapped. These lines represent the boundary between the flooded water and the dry land, therefore it is also referred to as the Water-Land Boundary. The SAR-based extent and LiDAR lines are assessed by visual comparison.



(a) The selected LMW locations.

(b) Water Depth Model computed by Deltares.

(c) Aerial photograph.

(d) Water-Land Boundary lines.

Figure 2.17: The four validation sources for the comparison of the SAR results.

2.8.5. Maximum flood extents computed for the ENW-report

The Dutch Expertise Network for Flood Protection (ENW) published a report with the preliminary explorations on the floods in September 2021. The maximum flood extents for this event have been determined by available modeled data with additional ancillary data from local knowledge and photographs. These sources combined yielded the best possible determined maximum flood extent, at the time. The organizations that helped computing these results are, among others, Deltares, VU Amsterdam and TU Delft. These extents will be used to compute a general overview of the results, indicating how many pixels are inside the flooded areas. [8]

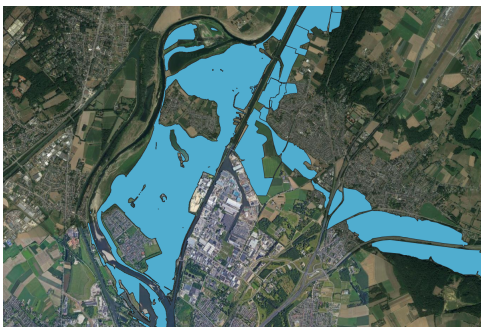


Figure 2.18: The computed maximum flood extents by the ENW report

3

Methodology

A few methods for SAR based flood mapping have been mentioned in section 2.5. Various methods for extracting the extent of the flood are explored and compared in this thesis. The final choices pursued will be partly based on performance and data availability. Furthermore, the methods are chosen in order for future research or future practical studies to be able to reproduce the results for other areas in the Netherlands that are subject to flooding. The methods should also represent the additional value contained in the SAR signal, as this is the main research goal of the thesis. The validation of the results will be against other resources that have been computed during the 2021 floods. This chapter will explain several researched methods for the different data sources and how validation will be performed. The methods are divided into multiple sections. First, water classification methods for the derivation of the flood extent, subsequently the computations of water levels are described. The last section is dedicated to the validation of the derived results.

3.1. Water Classification

3.1.1. Thresholding

Thresholding is one of the most widely used techniques for SAR-based flood mapping. The technique is based on the difference in backscatter values between water pixels and land pixels. It is a simple, yet efficient, approach to segmenting the image without the need for high processing power. Thresholding can be approached as a classic binary segmentation, it applies a threshold to separate flooded (dark) and non-flooded (bright) pixels [29]. Optimizing the threshold can be done by analyzing the backscatter signature profiles of flooded or water pixels [40].

Using the specular reflection phenomenon, water bodies can be distinguished from land pixels. Water pixels are expected to present low backscatter values because the signal is scattered away from the satellite. This principle is the basis for the most frequently used techniques for SAR flood mapping [29]. Thresholding can be performed on single images, image pairs, or a stack of images. Using an image pair or stack of images includes the concepts of the Change Detection approach. When using a single image, finding the most suitable threshold can be done by selecting a few reference areas. When computing a threshold range, one might even be able to distinct flooded areas from permanent water bodies because flood water is usually mixed with more debris. The presence of debris can increase the backscatter values, making possible to separate water pixels from flooded pixels [25]. Studies by Manjusree et al. have taken several reference areas of flooded pixels to evaluate the optimum thresholds for SAR data with different polarizations. The thresholds have been determined by looking into the backscatter signature profiles of the selected pixels [40]. The study concluded flood water threshold ranges for different polarization types: −8 to −12 dB, −15 to −24 dB, and −6 to −15 dB in HH, HV, and VV polarizations, respectively. The results showed that HH polarization can make a better distinction between land and water surfaces. VV polarization showed potential to identify flood water separate from partially submerged areas. The latter feature can be important when dealing with damage assessments. Using the optimized values of these ranges can save costly time when dealing

with near real-time computations. For this thesis, the optimum thresholds for HH and VV polarization will be tested. Since the studies have been performed on RADARSAT-2 data instead of data acquired by Capella Space or Sentinel-1, the ranges need to be adjusted. Mathematically, the thresholding method is implemented as shown below:

$$P = \begin{cases} P \leq r_l & \text{Flooded} \\ P \geq r_u & \text{Flooded} \\ r_l \leq P_d \leq r_u & \text{Not Flooded}, \end{cases} \tag{3.1}$$

where P is an image pixel to be classified, and r_l and r_u are the lower and upper ranges, respectively. The classified image is a binary map with the flooded and non-flooded pixels. Computing the classification with a single image is prone to outliers due to speckle noise and areas with similar backscatter values to water, such as runway roads. The advantage of this method is that it requires one image and is computationally inexpensive to perform. To decrease the number of outliers within the classified image, a similar thresholding approach can be applied to two images. In this case, thresholding is performed on the difference between an image pair (one pre-event and one co-event image) instead of one image. Computing the difference beforehand can help identify coherent regions with a constant spectral signature. A significant increase or decrease in backscatter is not expected for coherent regions. In contrast, flooded regions are expected to show a temporary decrease in backscatter relative to the pre-event image. The threshold set for the difference image reflects the magnitude of change for each pixel. A threshold of 2σ implies that flooded pixels in the co-event image have changed with 2σ with respect to the pre-event image. Taking a statistical value instead of a given threshold retains backscatter information (flood probability) instead of a binary flood map. Most pixels will show a decrease in backscatter, however, an increase in backscatter could also be indicative to flooding. Areas with high water levels are expected to drop in amplitude, however, partially submerged areas or urban areas can also increase in backscatter due to the increased double bounce effect [27]. The implementation of thresholding on a pre-event and co-event image pair is displayed in Equation 3.2 and Equation 3.3.

$$P_d = P_f - P_{nf}, \tag{3.2}$$

$$P = \begin{cases} P_d \leq T_l & \text{Flooded} \\ P_d \geq T_u & \text{Flooded} \\ T_l \leq P_d \leq T_u & \text{Not Flooded}, \end{cases} \tag{3.3}$$

where P_d is a pixel in the difference image, P is an classified image pixel, P_f is the co-event image for flooded conditions, P_{nf} is the pre-event image for normal conditions, T_l and T_u are the lower and upper threshold respectively. Similar to single image thresholding, this method is a simple technique with low computational costs. The increasing complexity of the second thresholding method is expected to result in fewer false detections due to stationarity within an area. Further improvements are taking the mean or median image values of a stack of pre-event images, to reduce the presence of speckle. The thresholding method is preferred for a simple analysis with low computational costs and utilizing a robust segmentation method. One of the weaknesses of the method is the relative extent of the flood compared to the full image and the contrast between flooded and non-flooded pixels. Thresholding does not take into account the local variations that arise. In addition, the contrast between flooded and non-flooded pixels becomes complex in areas with shadows or roads. Misclassfications can occur due to the similar backscatter profiles to water pixels. By including change detection and adopting a tiling approach over the image, local variations can be taken into account more accurately. However, the choice of pre-event or baseline image remains critical, as backscatter values continuously change over time. To identify the pixels that are specifically related to flooding, assumptions have to be made.

3.1.2. Otsu Thresholding

Otsu thresholding is an alternative, widely utilized, thresholding approach for SAR-based flood monitoring. The algorithm behind the method determines the optimal threshold automatically. It works by maximizing the interclass variance of two classes from a gray-level histogram by a nonparametric and unsupervised threshold selection [41]. Advantages of unsupervised methods are computational costs,

large-scale employment and the absence of ground truth data, being more efficient compared to supervised approaches. The method is applied to the intensity histogram of a pre-processed SAR-image, the algorithm computes the threshold defined as a weighted sum of the variances (σ^2 between the two classes in the image). The weights of the two classes are the probabilities (flooded: P_f and unflooded: P_{nf}) being separated by a threshold (t). Equation 3.4 shows the threshold estimation from the probabilities and variances of the two classes (flooded: σ_f and non-flooded: σ_{nf}). The threshold can be applied to a single image or to a stack of images (dynamic Otsu thresholding). The computed threshold is thereafter applied similarly to Equation 3.1, to classify the image. The equation of threshold is

$$\sigma^2(t) = P_f(t) \times \sigma_f^2(t) + P_{nf}(t) \times \sigma_{nf}^2(t). \tag{3.4}$$

The method can also be used to determine differences between the pre-event and co-event image. In this scenario, the dynamic Otsu threshold is applied to the pre-event image that generates a water / no-water map at time *t-1*. The co-event image is processes similarly using the Otsu threshold, providing a water / no-water map at time *t* [42]. Using both maps, a distinction can be made between flooded pixels and permanent water pixels by examining the state of the pixel in both maps:

1. If the state of a pixel at time *t-1* is no-water, and the state at *t* is water: the pixel is classified as flooded.
2. If the state of a pixel at time *t-1* is no-water or water, and the state at *t* is no-water: the pixel is classified as not flooded.
3. If the state of a pixel at time *t-1* is water, and the state at *t* is water: the pixel is classified as flooded (permanent water).

The Otsu method is a simple, efficient and widely used thresholding approach for SAR-based flood mapping. However, the complex characteristics of a SAR image can make the approach challenging to use for real-life applications. The contrast between pixels due to the presence of emergent vegetation, debris or waves could influence the spread in backscatter, thus the results of the threshold estimation. The Otsu algorithm does not take into account these local effects but are related to flooding. For this reason, the earlier explained approach of backscatter signature profiles will be applied in this thesis. By computing the backscatter profiles for several selected areas, the threshold results can be directly linked to the spread of backscatter values found for flooded pixels. Presence of emergent vegetation or other influences can be evaluated to generate a threshold that is suitable for the SAR image.

3.1.3. Time Series Analysis

Flood monitoring can be further refined by using time series. Exploiting the multi-temporal features of the SAR data can identify the dynamic changes observed during flood events. The Time Series Analysis (TSA) approach estimates the normal backscatter conditions for each pixel over time. Each pixel is assumed to follow a trend over time. Flood events are expected to show in the time series as outliers or spikes that can be detected. TSA is a more complex approach compared to the thresholding but has been proven to yield more accurate results [43].

Bayesian Framework

Recent methodologies propose to use the SAR time series in combination with modeled probabilities, for example by adopting a Bayesian framework [31]. Bayesian frameworks have been widely used for floods, regarding: monitoring, forecasting, detections, modeling and mapping. The method involves an estimation of posterior probabilities over an area that can be conditioned by observations through the application of the Bayes equation. This equation is a function of the probabilities of the presence or absence of a flood event at each observation (also referred to as, the likelihood of a flood event). To estimate the likelihoods, ancillary data such as terrain height or the Height Above Nearest Drainage can be used. The presence of flooded water (*F*) inside a pixel (*P*) at a specific timestamp (*t*) is then modeled as a conditional probability ($p_{P,t}(F|\sigma^0)$ or simple notation: $P(F|\sigma^0)$).

$$p(F|\sigma^0) = \frac{p(\sigma^0|F)p(F)}{p(\sigma^0|F)p(F) + p(\sigma^0|\overline{F})p(\overline{F})}, \tag{3.5}$$

where $p(\sigma^0|F)$ and $p(\sigma^0|\overline{F})$ are the likelihoods (or probabilities) of the observed backscatter, conditioned by the presence or absence of a flooding event, respectively. $p(F)$ and $p(\overline{F})$ indicate the prior probability of flood or no flood, respectively, based on ancillary data [31]. Results of a Bayesian framework are flood probability maps that contain a pixel-based confidence about the flooded conditions for each date in the time series.

There are a few shortcomings of this method when it comes to the application to flood events in the Netherlands. The estimations of the prior probabilities requires knowledge about the terrain and how risks of the area are identified. Flooding in the Netherlands is not as common as in other European countries such as Italy. Proper risk assessments in the areas are complex to approach. Due to the lack of data of previous events, training and computation of statistical inference mainly relies on assumptions. More training data is needed for any machine learning approaches to accurately perform. Acquiring more data in the future could lead to more advanced modeling methods. For this thesis, focus will be on a time series approach that is based on available SAR data while limiting the amount of assumptions to be made. In this manner, the results will be derived from the information contained within the SAR signal, rather than being based on assumptions.

Amplitude Time Series Analysis

TSA can be performed without using posterior knowledge or ancillary data but solely focusing on the variations of backscatter over time. Sentinel-1 data has the advantage of having a large data archive, making it being favorable for TSA methods. The Sensar's Amplitude Time Series Analysis (ATSA) algorithm is usually applied to InSAR applications, however, it can be optimized for outlier removal concerning temporal events such as floods. Focus of the algorithm is on multi-temporal variations of the amplitude values and applying outlier removal. Outlier removal is performed based on determined significance values, which are derived from the dispersion of the data. The full explanation of the method is described in Appendix A [1]. Detected outliers of the algorithm can be related to events such as maintenance, harbor activities or flood events. The settings of the method can be optimized for flood detection and generate an outlier mask for the image acquired during a flood event. Several optimizations can be implemented to this approach such as despeckling of the data to generate smoother time series and data dispersion.

3.1.4. Coherence and Polarization

Thresholding methods and TSA approaches can be performed on amplitude and coherence data. Within both components a choice can be made between different types of polarization, if available for the chosen data product. Sentinel-1 is an example that offers both coherence data and different polarization. The use of coherence data can provide additional information to amplitude data on locations where the SAR signal is complex, such as urban areas. A downside of using coherence data is the decrease in resolution which is not preferable when comparing Sentinel-1 to a sub-meter resolution source such as Capella Space. Other complications of coherence are the presence of vegetation, were coherence is generally low throughout the entire time series. Detecting coherence changes related the presence of floods is mostly of interest in high coherent areas such as cities [30].
Recent studies by Pelich et al. have shown the improvements of urban flood mapping by using both co- and cross-polarization channels [27]. The implementation of multiple polarizations can improve a flood map based on more complete and full exploitation of the SAR signal. Generally, VV polarization is more sensitive to vertical structures and performs better for open water bodies or detecting the edges of vegetation. VV is known to yield more accurate flood mapping results and is therefore a more widely used polarization choice. VH polarization can perform better over terrains with more roughness. The roughness of the terrain can affect the VV polarized signal significantly, making the extracting of the flooded area more accurately using VH polarization [44]. A downside of using multiple polarizations is a decrease in spatial resolution due to signal averaging and an increased computational load.
As will be further explained in the next section, the spatial resolution of the data is important in this research regarding the computation of water levels. Accurate extraction of the water levels is highly dependent on the flood extent. Considering that both coherence and polarimetry decrease the spatial resolution of the data, only amplitude data is considered for this thesis.

[1] Appendix A contains confidential information and is left unpublished.

3.2. Water Level Estimation

To compute the water levels of the flooded area, both longitudinal and transverse gradients are taken into account. Rather than assigning a single value to each flooded polygon, the outer edges of the polygon are sampled. This approach eliminates the assumption that water levels inside a flooded polygon are at constant elevation. The edges of the polygons are sampled every 10 meters, and the corresponding DEM value is assigned to all sampled locations. In this way, the results reflect the variations in water levels along the boundaries. An example of the sampling is given in Figure 3.1. For each sampled location the following values from the DEMs are extracted:

- DTM height [m]: AHN4-DTM or DHMV2-DTM
- DTM height variations [m/10m]: Standard Deviation AHN4-DTM within 10 m or Standard Deviation DHMV2-DTM within 10 m
- DTM Slope [degrees]: AHN4-DTM or DHMV2-DTM
- DSM height [m]: AHN4-DSM or DHMV2-DSM
- DSM height variations [m/10m]: Standard Deviation AHN4-DSM within 10 m or Standard Deviation DHMV2-DSM within 10 m
- DSM Slope [degrees]: AHN4-DSM or DHMV2-DSM

The above-mentioned parameters entail information about the water level of all the edge points, plus some quality metrics to provide how much variation and steepness there is on the locations. In order to further express the quality of the estimated water levels, the precision is provided over the sampled locations. The standard deviation on each sampled location is a measure of the water level precision. It is computed under the assumption that all edge points are independent measurement points. The standard deviation is determined by computing the average difference in water level ($\overline{\Delta WL}$) between two neighboring points, resulting in a standard deviation (σ) for a single point of

$$\sigma = \overline{\Delta WL}/\sqrt{2}. \tag{3.6}$$

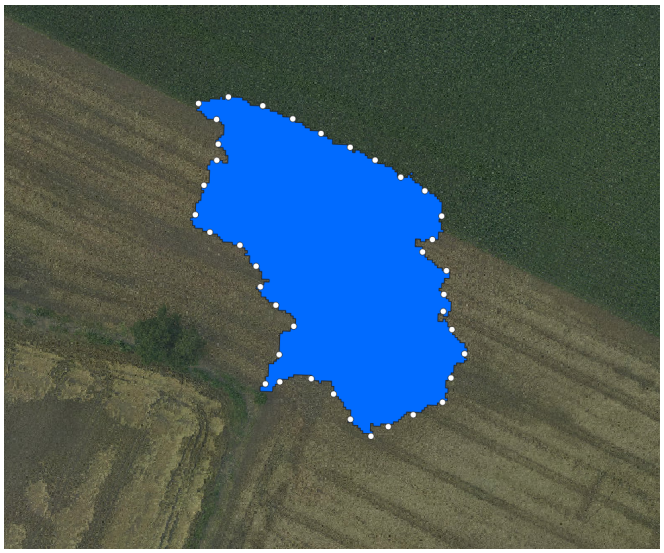


Figure 3.1: Overview of sampling the boundaries of the flooded polygon every 10 meters, over which the DEM values will be computed to extract the water levels.

The accuracy of the estimated water levels is directly linked to the definition of the flood extent. In Figure 3.2, the relation between the (x,y) estimate and the computation of the water level (z) is displayed. Over- and underestimation of the flood extent is therefore related to over- and underestimation of the water levels.

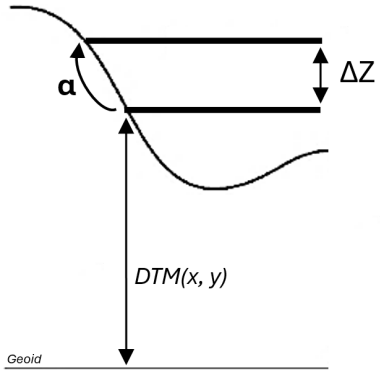


Figure 3.2: The influence of a changing (x,y) location of the flood extent on the height estimate (z), on a DTM with a slope α.

As can be seen from the figure, the angle of the slope can influence the amount of variation of height within an area. The spatial resolution of Capella Space imagery is less than 1 m, therefore one misclassified pixel has a little effect. The spatial resolution of Sentinel-1 is lower, one or two misclassified pixels can lead to large differences in water level. Interferences that cause underestimation of the extent can have significant effects on the derived water levels as well. Based on averaged results within the AOI of the case study, an indication of slope effects is provided below:

- Slope 0-5 degrees: up to 50 cm difference in height within 10 m radius.
- Slope 5-10 degrees: 50 cm to 90 cm difference in height within 10 m radius.
- Slope > 10 degrees: 90 cm to 2 m difference in height within 10 m radius.

3.3. Accuracy Assessment

3.3.1. Validation methods

The results of the thesis will be described in the next chapter. The results are evaluated with respect to other resources, to access the accuracy of the flood extent. For the flood extent, an error matrix will be employed. Water level results are evaluated to other resources by generating a frequency distribution and statistics in close vicinity to the gauges.

Error Matrix

For classification studies, a conventional method to validate results is a Confusion Matrix, or also referred to as Error Matrix when dealing with two datasets that are subject to errors and misclassification, (Figure 3.3). To construct an error matrix, one dataset will be designated as the reference, and the second dataset is compared to it. The matrix is designed to show the misclassifications of either of the two datasets. The error matrix displays the amount of measured (SAR-based) and modeled values, grouped by: True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). True Positives and True Negatives represent a correct, or consistent, classification of flooded pixels of the two datasets. True Positives refer to the flooded-classification of a pixel: both measured and modeled pixels are classified as flooded. On the contrary, when both datasets classify a pixel as non-flooded is labeled as a True Negative. In this case neither of the datasets detected a flooded state. False Positives are a Type I error, it indicates that the modeled class did not classify the pixel as flooded but the measured dataset did. Vice versa holds for the False Negatives, these are type II errors that refer to the modeled class being flooded but the measured pixel is not. The error matrix can be used to compute several performance parameters:

- Accuracy, represents the amount of correct classifications to the total number of classifications. It is a measure of performance to predict the correct class by the measured data, this variable can be misleading in imbalanced datasets. For example, where the amount of non-flooded pixels is much higher compared to the amount of flooded pixels:

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)}. \tag{3.7}$$

- Precision, denotes the ratio of True Positives to the total number of Positive classifications. This parameter indicates how trustworthy the positive predictions are. A lower precision indicates a higher amount of False Positives, which can be related to overestimation of the flood by the measured dataset or misclassifications of the modeled dataset:

$$\text{Precision} = \frac{TP}{(TP + FP)}. \tag{3.8}$$

- Recall (Sensitivity), evaluates the amount of True Positives to the total number of actual Positives. The Recall gives an indication about the correct classifications of the measured dataset with respect to the modeled dataset. This parameter is important to evaluate for the amount of consistent classified pixels between the datasets. A lower recall indicates an underestimation of the measured dataset with respect to the modeled dataset or an overestimation of the modeled data:

$$\text{Recall} = \frac{TP}{(TP + FN)}. \tag{3.9}$$

- Specificity, evaluates the amount of True Negatives to the total number of actual Negatives. This parameter is important for classification of non-flooded areas. The specificity can be used to describe the amount of overestimation of flooded pixels from the measured data if the amount of False Positives is high:

$$\text{Specificity} = \frac{TN}{(TN + FP)}. \tag{3.10}$$

- F1-Score, a combination parameter that balances the precision and recall. This is a good parameter when both False Positives and False Negatives are important:

$$F_1 = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}. \tag{3.11}$$

All the parameters are also schematically displayed in the figure below:

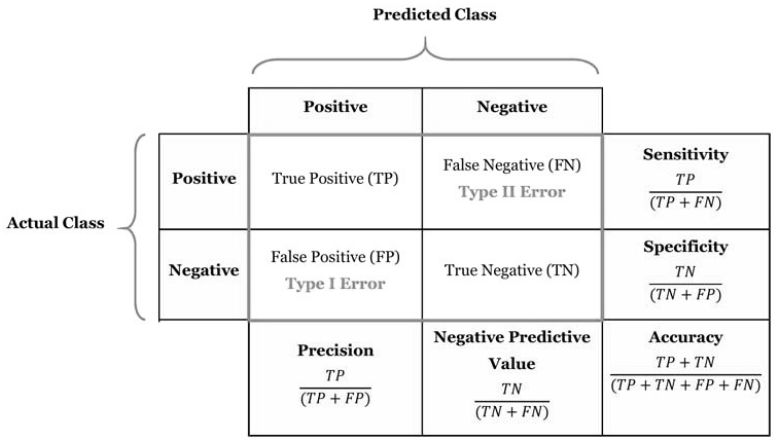


Figure 3.3: Overview of a Confusion Matrix or Error Matrix with the performance parameters [45]. The Actual Class and Predicted Class refer to the Modeled Class and Measured Class, respectively.

Frequency Distribution

Frequency distributions are generated to look into the offsets between estimated water levels from SAR-based flood extents compared to the measured and modeled water levels. Frequency distributions can be used to show the absolute or relative frequencies of a chosen dataset. The values can be presented in the form of a table or graphs, this choice is dependent on the data type. For this thesis, histogram charts the most suitable option. Histograms can show the distribution of all the observations in a dataset. By examining the shape and spread, the most dominant values can be determined. Performing this analysis for the water level results can provide insights into how well the estimated data correlates with the validation data.

Water level differences

Differences in water level (WL_{dif}) between the SAR-based water levels (WL_{SAR}) and the modeled results (WL_{model}) can be computed with

$$WL_{dif} = WL_{\text{model}} - WL_{\text{SAR}}. \tag{3.12}$$

The results are projected to all the sampled points on the edges of the flood extent polygons. Visual comparison and frequency distributions will be generated to evaluate the differences between the datasets. In case WL_{dif} is negative, this indicates that there is an overestimation of the SAR-based water levels. Vice versa, if WL_{dif} is positive it means an underestimation of the SAR-based water levels with respect to the water level results from the model data. The accuracy of the water levels required by Rijkswaterstaat is 20 cm.

4

Results

This chapter will display the results of the methods explained in chapter 3. The results of the water classification methods will be presented first, followed by the water level estimates. Both sections present the results separately for the Capella Space and Sentinel-1 data. For all methods, a brief explanation will be given on further improvements of the results. The flood extent will be shown in a blue or off-white color that is displayed on top of a map. The comparison of the SAR results with aerial images and LiDAR data will be discussed in the next chapter. The results chapter will aim to focus on the comparisons with the flood extents and water levels derived from the hydrodynamic model from Deltares and the LMW-gauges from Rijkswaterstaat. For all data sources, the computations have been performed over the same AOI. By fixing the AOI, the differences can be fairly compared in chapter 5. The chosen AOI is an area that was severely affected during the floods in 2021. Therefore, this case study can provide insights into the potential of real-life application of the chosen data types. The used images of Sentinel-1 and Capella Space during the flood event on the 15^{th} of July 2021 are displayed in Figure 2.14.

4.1. Capella Space - Flood extent

The sub-meter resolution images from Capella Space have been pre-processed before extracting the flood extent. For this thesis, the thresholding approach is used and is further explained in chapter 3. The thresholding technique has been implemented in two ways: on a single co-event image and on an image pair of a co-event and post-event image. In general, thresholding on a pair of images is applied on a pre-event and co-event image. There was no pre-event image of the area acquired by Capella Space, so the post-event is taken to analyze the additional value of thresholding of an image pair.

4.1.1. Thresholding - Single Co-event image

To determine the optimum threshold ranges for water pixels, several regions have been spatially distributed within the AOI. The regions are used to evaluate the backscatter profiles on permanent water, flooded land with low vegetation and flooded land with high vegetation. The regions are shown on top of the SAR image acquired by Capella Space in Figure 4.2. The statistics of the backscatter profiles from the selected regions are presented in Table 4.1.

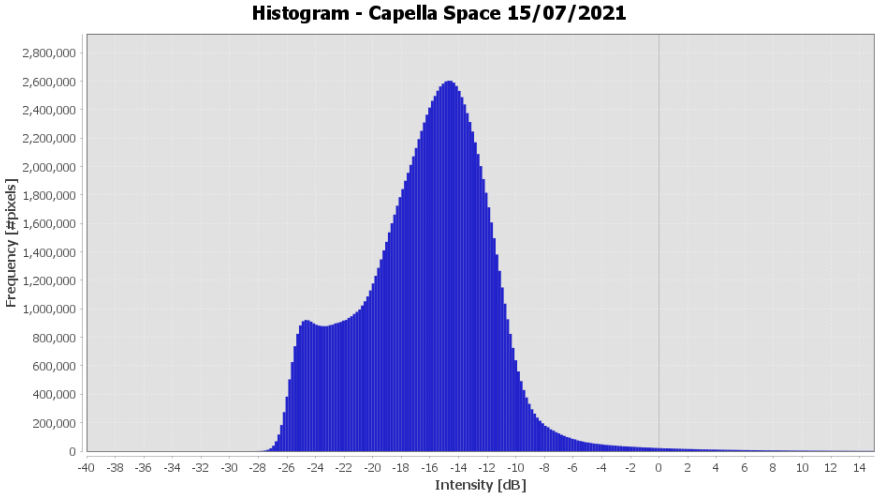


Figure 4.1: Histogram of the entire image of Capella Space acquired on 15/07/2021.

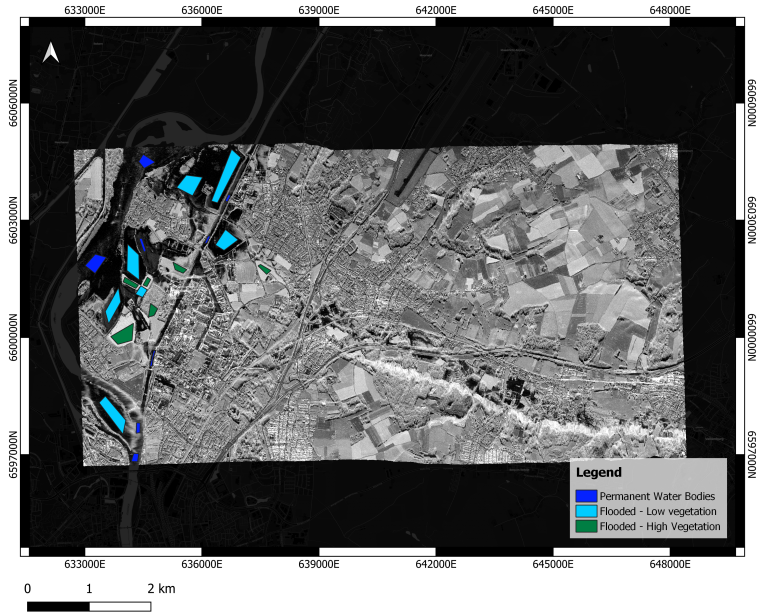


Figure 4.2: Regions used to evaluate the backscatter of flood water.

Table 4.1: Backscatter values related to the selected regions of the SAR image.

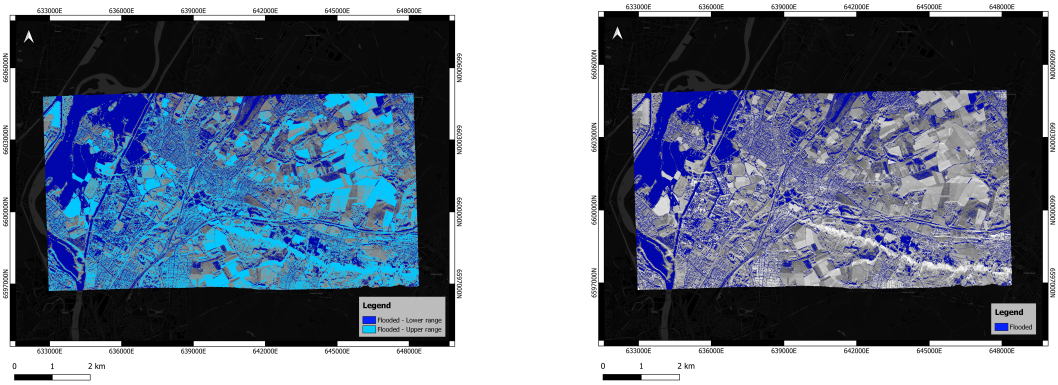
Region	Median [dB]	Minimum [dB]	Maximum [dB]
Permanent Water	-23.15	-27.27	-19.08
Flooded - Low vegetation	-24.75	-28.62	-15.34
Flooded - High vegetation	-10.91	-14.36	-3.47

From Table 4.1 and Figure 4.1 it can be derived that the lower backscatter values of water are among the lowest values in the image, which is expected due to specular reflection. Based on the results of the regions and histogram, the thresholding ranges have been determined and applied using Equation 3.1. The equation for the thresholding that results from implementation is

$$P = \begin{cases} P \leq -19\ dB & \text{Flooded} \\ P \geq -14\ dB & \text{Flooded} \\ \text{Otherwise} & \text{Not Flooded.} \end{cases} \tag{4.1}$$

The results of the classified pixels are shown in Figure 4.3a. From Table 4.1, it can already been seen that flooded areas show a larger spread in backscatter values compared to permanent water. These effects are due to emergent vegetation and debris that arise with the dynamics of a flood. Densely vegetated areas also show a large spread in backscatter due to the increased double bounce effects. This latter effect is visible in all submerged agricultural areas and is displayed in Figure 4.3a. The classifications include fields that are far beyond the edges of the maximum flood extent of the ENW report, indicating the presence of false positive detections. To reduce the number of false positives, only the ranges on low vegetated areas are considered, obtaining the results of (Figure 4.3b). Inserting these decisions into Equation 3.1 yields the following final matrix for the co-event image threshold:

$$P = \begin{cases} P \leq -19\ dB & \text{Flooded} \\ \text{Otherwise} & \text{Not Flooded.} \end{cases} \tag{4.2}$$



(a) Thresholding with both lower range and upper range.

(b) Thresholding with the lower range.

Figure 4.3: Capella Space co-event image thresholding with different ranges.

The resulting classifications of the applied threshold are shown in Figure 4.3b. Many of the water bodies and flooded areas are correctly classified when compared to the maximum flood extent. However, the results also include false classifications such as puddles, shadowing and roads that have a similar backscatter footprint as flood water. To limit the influences of these classification into the flooded polygons, the classified pixels are clustered and filtered on areas that are larger than 2500 m^2 (50x50 meters). Water levels are considered to be constant within a radius of 50 meters, so this is the minimum area that is considered as a flooded polygon. The final flood extent is shown in Figure 4.4. Compared to the maximum extent of the ENW report: 41% of the flooded polygons are permanent water, 17% is inside the maximum flood extent and 42% is outside the maximum flood extent. These numbers indicate that this method is sensitive to false positives.

The final flood extent is clipped to the AOI and shown in Figure 4.5a. The extent is compared to the results of the Deltares model using a composite. The consistent classifications between the two datasets are shown in Figure 4.5b, where the regions of overestimation or underestimation can be visually detected.

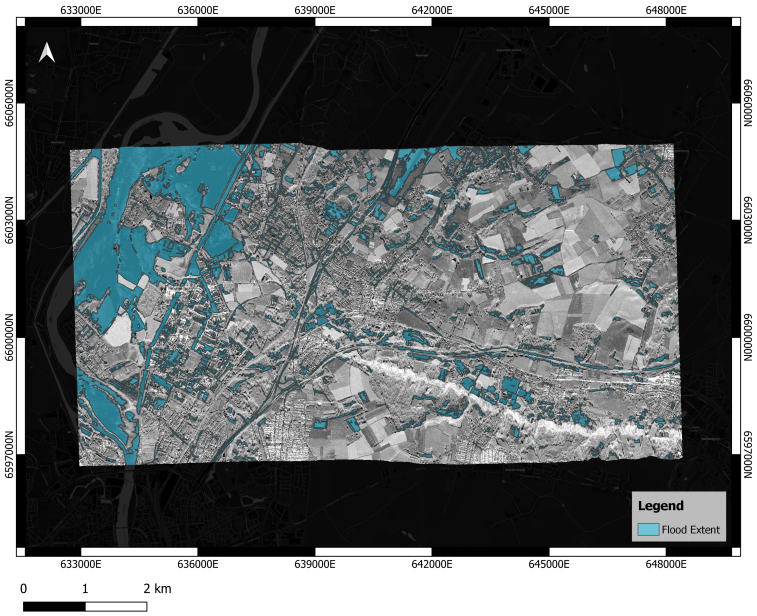
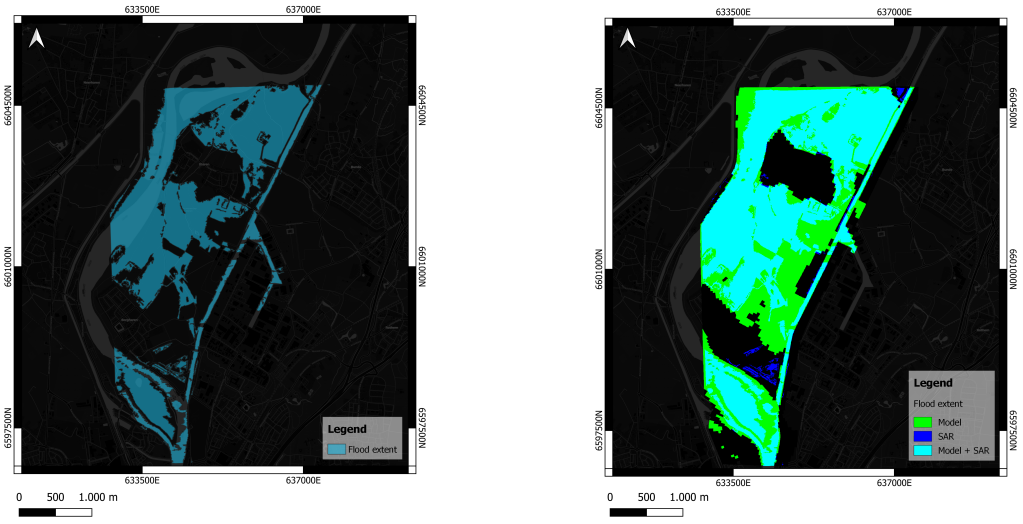


Figure 4.4: Final flood extent polygons derived from the Capella Space acquisition on the 15^{th} of July.



(a) Flood extent.

(b) Difference between SAR-based and modeled flood extent.

Figure 4.5: Final flood extent polygons derived from the Capella Space acquisition on the 15^{th} of July, considering only the flooded polygons relevant for flooding of the Meuse river.

To quantitatively compare the SAR-based flood extent to the modeled results, an error matrix is made for further evaluation. The error matrix displays the results of the SAR-based flood extent compared to the flood extent that the model shows at the local time 16^{th} July 00:30 CET (or 15^{th} July 23:30 UTC). The performance parameters of the error matrix are displayed in Table 4.2. The parameters show that the dataset is imbalanced between flooded and non-flooded pixels, making the accuracy of less interest. The precision and specificity are 0.92 and 0.87 respectively, indicating a low number of False Positives or low amount of overestimation of the SAR-based results. The recall is 0.67, indicating underestimation of the SAR-based results compared to the modeled data. The combined F1-score is 0.78, the areas of overestimation and underestimation of the SAR-based flood extent can be seen in Figure 4.6.

Table 4.2: Performance parameters of the error matrix between model results at 23:30 UTC and SAR-based flood extent of Capella Space.

Accuracy	Precision	Recall	Specificity	F1-Score
0.73	0.92	0.67	0.87	0.78

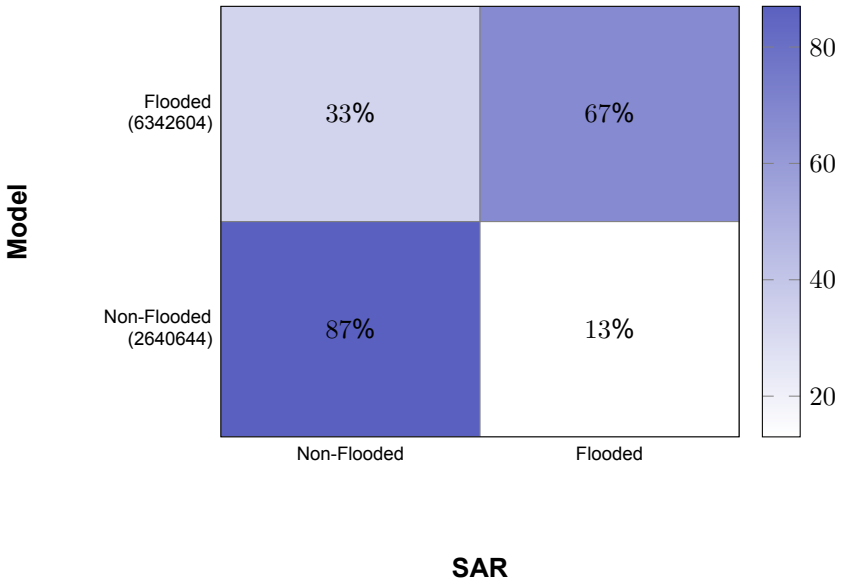


Figure 4.6: Error matrix of flooded and non-flooded pixels, between Deltares model and the SAR-based flood extent from Capella Space. The total amount of classified pixels are provided behind each class of the modeled results.

4.1.2. Thresholding - Co- and Post-event image pair

Applying a threshold to a single image includes the benefit of classifying all water bodies inside the image. The disadvantage is the amount of false positives caused by shadows or roads. To reduce the amount of false positives, an additional image can be used as a baseline to compute the differences between the images. For this thesis, only a co-event and post-event image were available. The post-event image does not cover the entire area covered by the co-event image, however, it is sufficient for demonstrating the use of a baseline image.

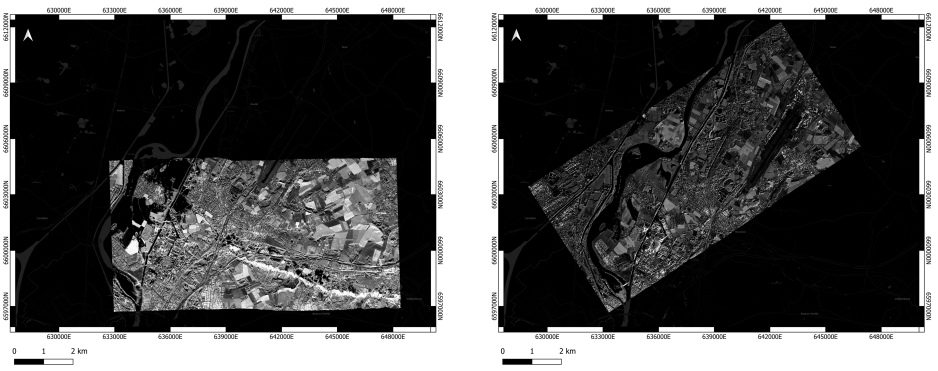


Figure 4.7: Acquisitions of Capella Space on the 15^{th} and 17^{th} of July 2021. © 2023 Capella Space. All Rights Reserved [12].

The baseline image of normal conditions will be represented by the image that was acquired two days after the peak of the flood. Differences greater than the standard deviation (σ) of the baseline image

are classified as flooded. Inserting this into Equation 3.3 yields the equations below for the co-event and post-event image pair thresholds

$$P_d = P_{co} - P_{post}, \tag{4.3}$$

$$P_d = \begin{cases} P_d \leq -1\sigma\ dB & \text{Flooded} \\ P_d \geq 1\sigma\ dB & \text{Flooded} \\ \text{Otherwise} & \text{Not Flooded.} \end{cases} \tag{4.4}$$

The results of the applied thresholds are visualized on top of the SAR image acquired on the 15^{th} of July, in Figure 4.8. False positive detections are reduced by using the baseline image, especially around areas such as airport runways. Both thresholding methods are compared in Figure 4.9, where the effects of a baseline can be seen. The classifications of permanent water bodies are also reduced. To obtain the full flood extent for this image, the permanent water bodies have to be detected separately. For this case study, it is assumed that limited data (one image) is available for the area. Companies that provide on-demand data often lack a historical database, although this may change in the future. The classification with a baseline image is solely performed to indicate the effects of change detection on the thresholding method with a baseline image. For this reason, no further computations or quality metrics are obtained for this method and it will be further discussed in chapter 6.

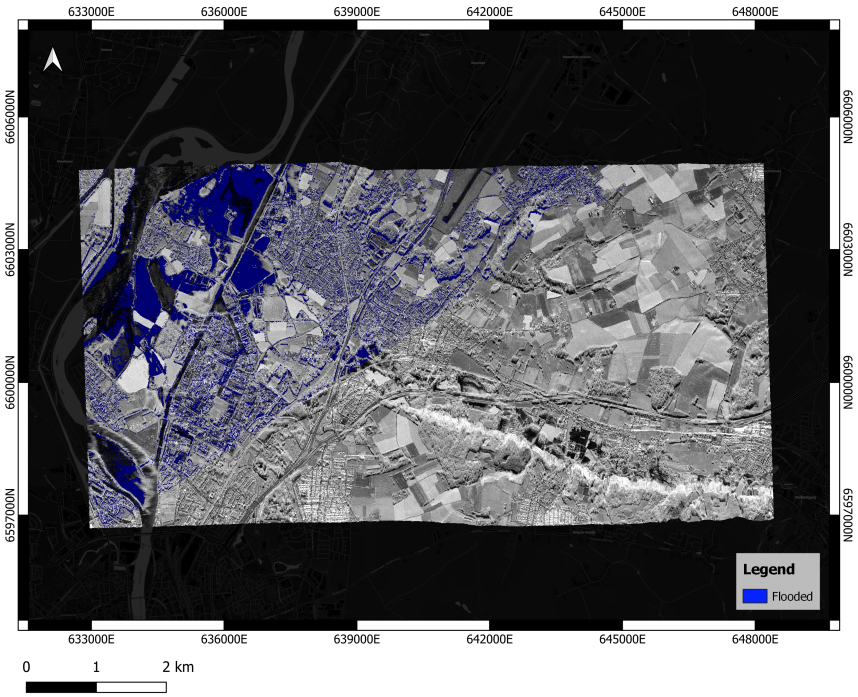
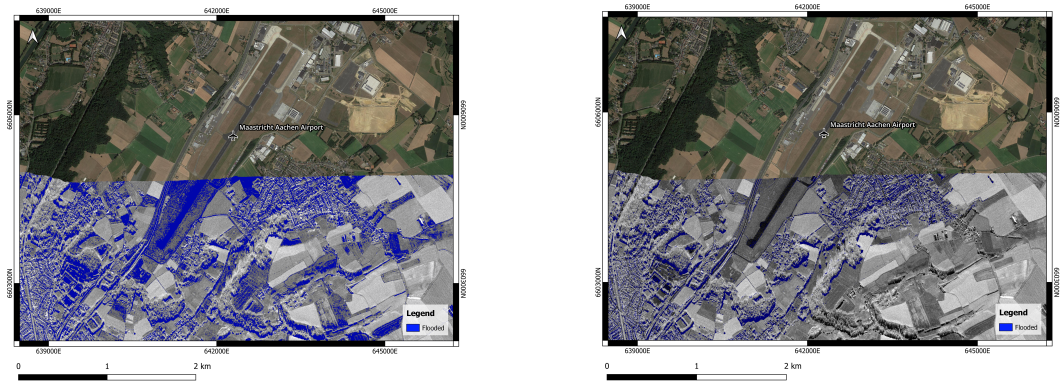


Figure 4.8: Thresholding the Co- and Post-event image with Equation 4.3 and Equation 4.4, acquired by Capella Space on the 15^{th} and 17^{th}, respectively.



(a) Co-event image thresholding results.

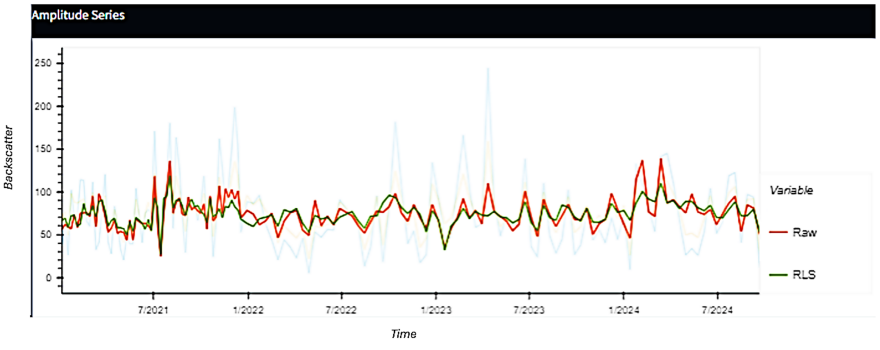
(b) Co- and Post-event image tresholding results.

Figure 4.9: Differences of classification around airports using only the co-event image or using an additional baseline image, displayed on top of the SAR and optical images of the area.

4.2. Sentinel-1 - Flood extent

The flood extent derived from Sentinel-1 data is computed in an alternative way compared to flood extent from the Capella Space data. As explained in chapter 3, full stacks with archives dating back to 2015 have been used to generate an amplitude time series for each individual pixel. The time series have been generated with both GRD and SLC data, displayed in Figure 4.10.

(a) GRD-data.



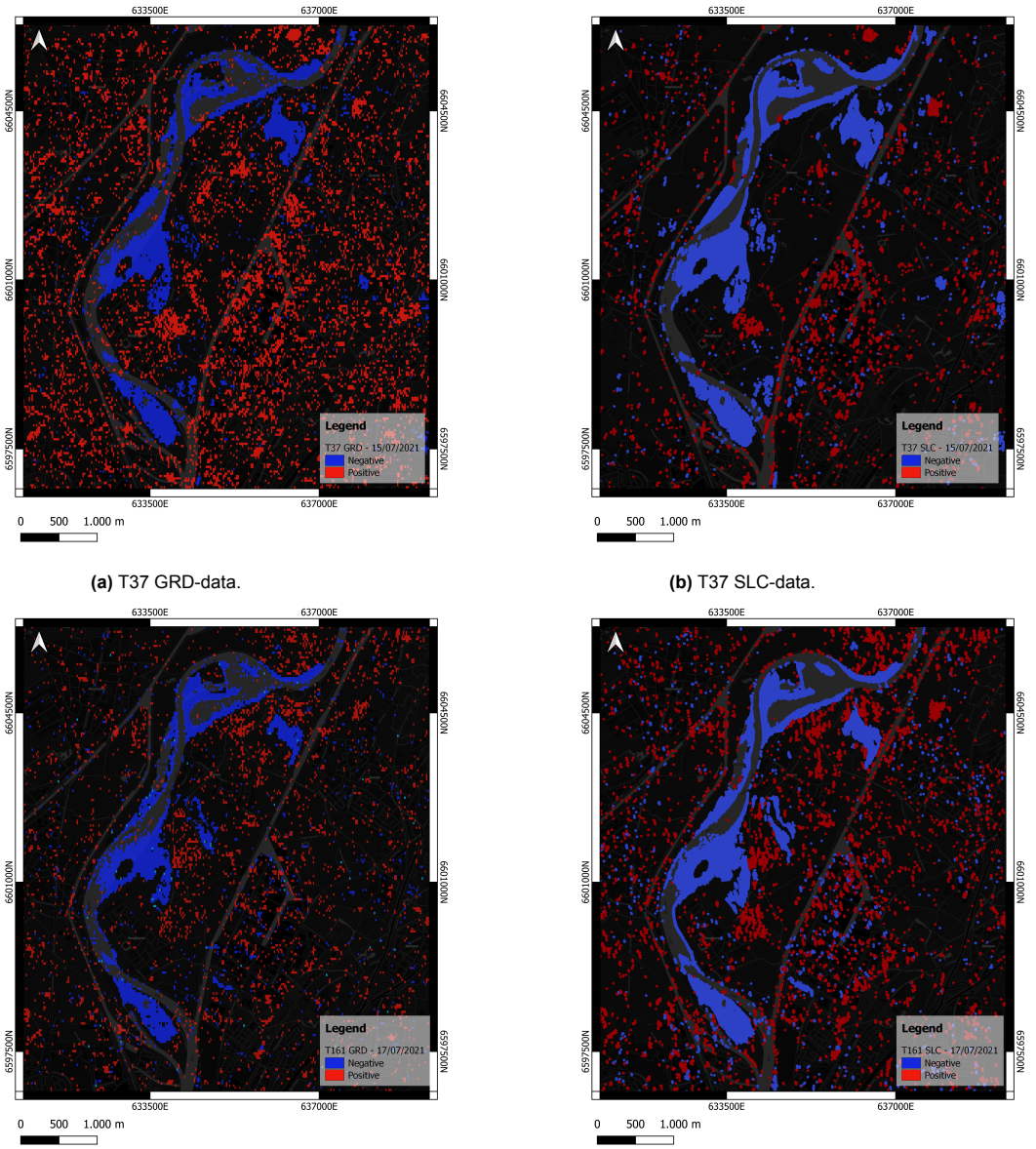
(b) SLC-data.

Figure 4.10: Time series of flooded pixels during the flood, for the Raw data (Red) and Refined Lee data (Green, RLS). On the acquisition on 15^{th} of July, a negative outlier can be seen.

As mentioned in section 2.7, there are different stacks available over the area. The Amplitude Time Series Analysis (ATSA) performs an outlier detection on the individual stacks. For each stack, the flood

extent can be derived from the corresponding outliers detected on that specific day. The first stack is T37, which acquired an image on 15^{th} of July at 05:50 UTC. This acquisition was right before the peak of the flood. The second stack is T161, which acquired an image on 17^{th} of July at 17:30 UTC, two days after the peak of the flood.

Using the ATSA method, outliers were detected on each date for the two different stacks. The results of SLC and GRD data are visualized in Figure 4.11. Blue detections refer to negative outliers (detected decrease in backscatter) and red detections to positive outliers (detected increase in backscatter). The percentages of outliers that are within the maximum flood extents of the ENW-report are summarized in Table 4.3. Based on the results in Table 4.3 the SLC data generates 10 to 20 percent more correctly classified detections. In addition, SLC data preserves higher spatial resolution than GRD data which is preferred for the extraction of water levels. Taking both variables into account, SLC data is the most appropriate selection in this case study.



(a) T37 GRD-data.

(b) T37 SLC-data.

(c) T161 GRD-data.

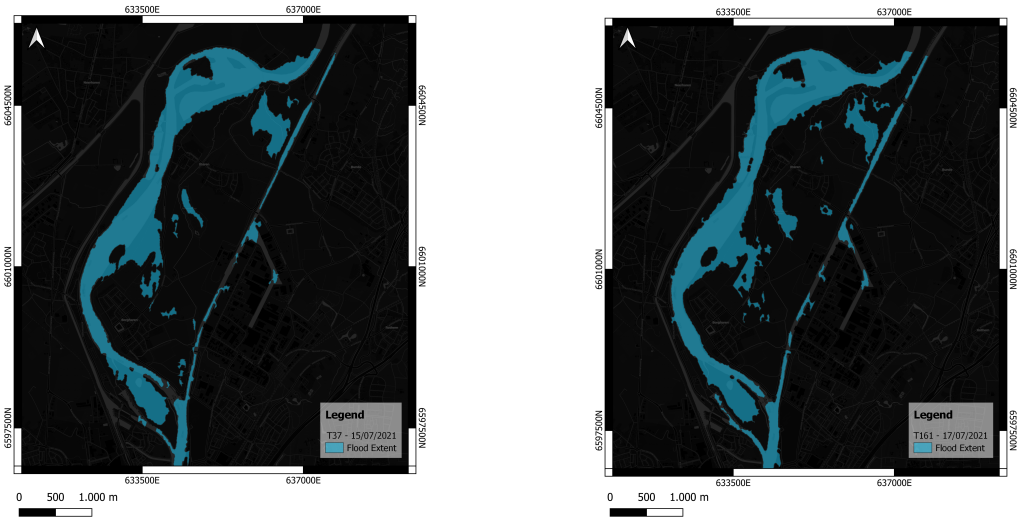
(d) T161 SLC-data.

Figure 4.11: Results of the ATSA detections, the pixel outliers detected on 15/07/2021 and 17/07/2021, for the T37 and T161 stacks respectively.

Table 4.3: Statistics of outliers detected by the ATSA approach within the maximum flood extent computed by the ENW report.

Stack	Data Type	Number of outliers	% Inside maximum extent
T37	GRD	11583 pixels	31%
T37	SLC	30282 points	51%
T161	GRD	7626 pixels	39%
T161	SLC	34799 points	51%

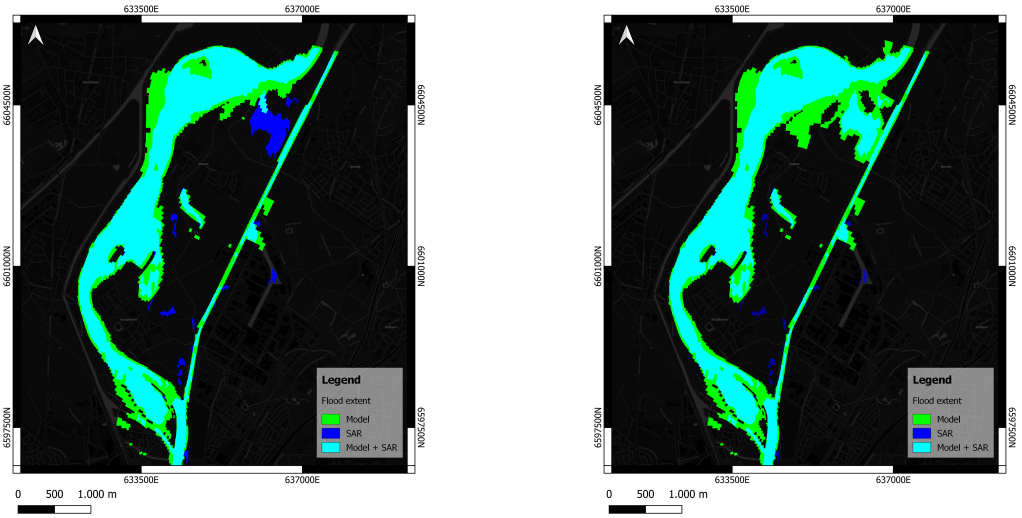
The detected outliers computed with the ATSA approach on SLC data have been clustered and filtered based on the size. Similar to the flood extent computed from Capella Space data, clusters with an area smaller than 2500 m^2 (50mx50m) have been filtered out. To generate the full flood extent, permanent water pixels were detected from the baseline layer. Combining the outliers with the permanent water extracted from the baseline layer yields the flooded polygons from Sentinel-1 on 15/07/2021 and 17/07/2021 as visualized in Figure 4.12. The combined flood inundation extent of both dates is shown in Figure 4.14. To quantitatively compare the SAR-based flood extent to the modeled results, an error matrix is made for the detections of the T37 stack on the 15^{th} of July compared to the modeled results at 5:50 UTC. The performance parameters of the error matrix are shown in Table 4.4. The time stamp of 05:50 UTC represents time of acquisition of the satellite, however, no significant changes are seen in the model simulations at this time. Therefore, the timestamp of 07:00 UTC is taken into account, results at this time stamp are visually in higher agreement with the flood extent detected by the SAR data. The performance parameters of the error matrix are displayed in Table 4.5
Additionally, an error matrix is made to compare the combined results of the flood inundation extent with the results of the model at 23:30 UTC, similar to the Capella Space data. This error matrix is made to observe how much of the total inundated extent can be detected by the combined Sentinel-1 acquisitions. Performance parameters are shown in Table 4.6. For each comparison, the consistency between the SAR-based flood extent and the modeled results are shown in the composites displayed in Figure 4.13 and Figure 4.14b.



(a) T37 - 15^{th} of July 2021.

(b) T161 - 17^{th} of July 2021.

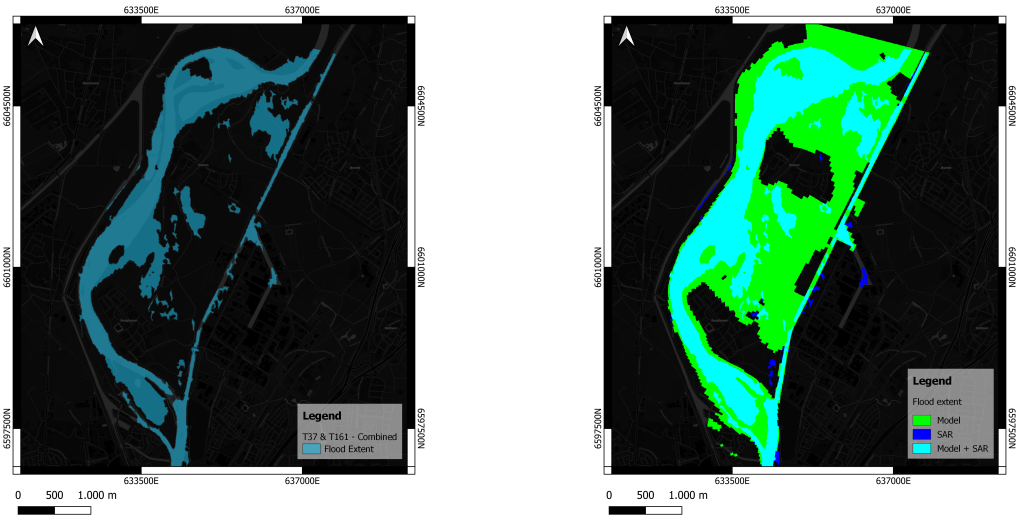
Figure 4.12: Final flood extent polygons derived from the Sentinel-1 ATSA computations.



(a) Model results at 05:50 UTC.

(b) Model results at 07:00 UTC.

Figure 4.13: Difference between SAR-based and modeled flood extent for the results from the 15^{th} of July 2021.



(a) Flood extent.

(b) Difference between SAR-based and modeled flood extent.

Figure 4.14: Final combined flood extent polygons derived from the Sentinel-1 ATSA computations on the 15^{th} and 17^{th} of July 2021.

Table 4.4: Performance parameters of the error matrix between the model results at 05:50 UTC and SAR-based flood extent of Sentinel-1 T37.

Accuracy	Precision	Recall	Specificity	F1-Score
0.94	0.82	0.68	0.98	0.74

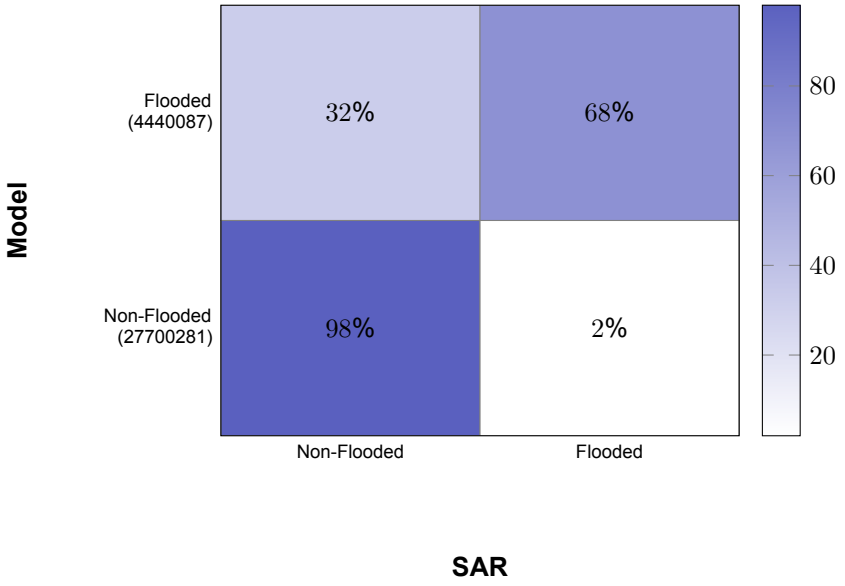


Figure 4.15: Error matrix of flooded and non-flooded pixels, between modeled results at 05:50 UTC and the SAR-based flood extent from Sentinel-1. The total amount of classified pixels are provided behind each class of the modeled results.

Table 4.5: Performance parameters of the error matrix between the model results at 07:00 UTC and SAR-based flood extent of Sentinel-1 T37.

Accuracy	Precision	Recall	Specificity	F1-Score
0.93	0.87	0.64	0.98	0.74

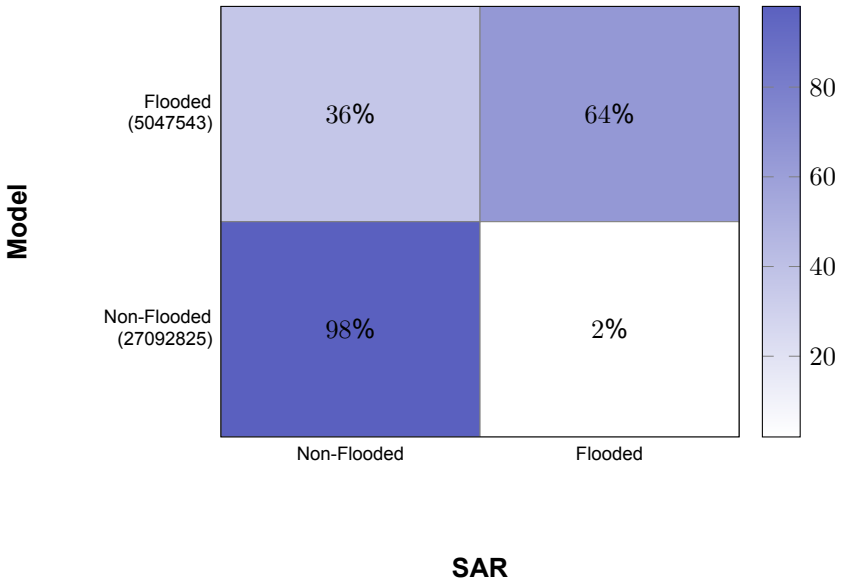


Figure 4.16: Error matrix of flooded and non-flooded pixels, between modeled results at 07:00 UTC and the SAR-based flood extent from Sentinel-1. The total amount of classified pixels are provided behind each class of the modeled results.

Table 4.6: Performance parameters of the error matrix between the model results at 23:30 UTC and SAR-based flood extent of Sentinel-1 T37 + T161.

Accuracy	Precision	Recall	Specificity	F1-Score
0.86	0.97	0.45	0.995	0.62

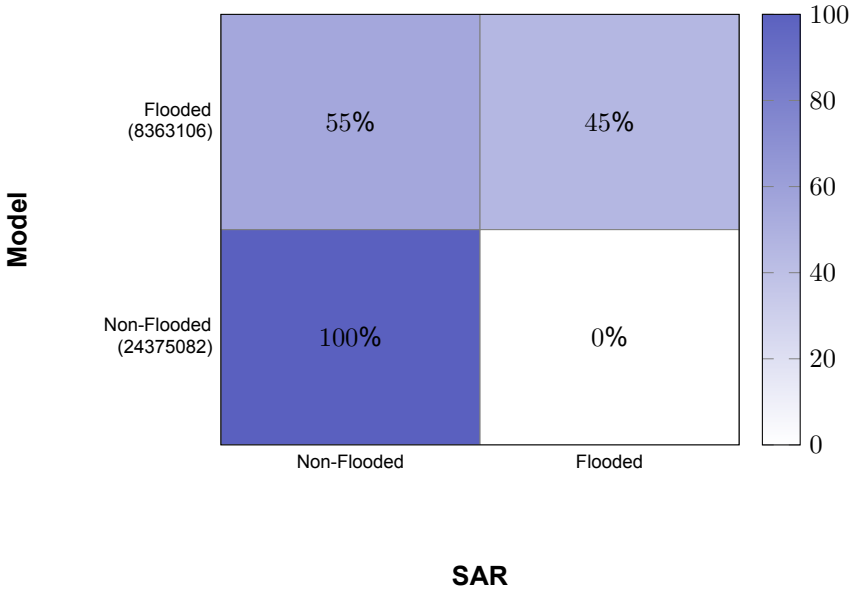


Figure 4.17: Error matrix of flooded and non-flooded pixels, between the results of the model at the peak of the flood and the SAR-based flood extent from Sentinel-1 combining both acquisitions. The total amount of classified pixels are provided behind each class of the modeled results.

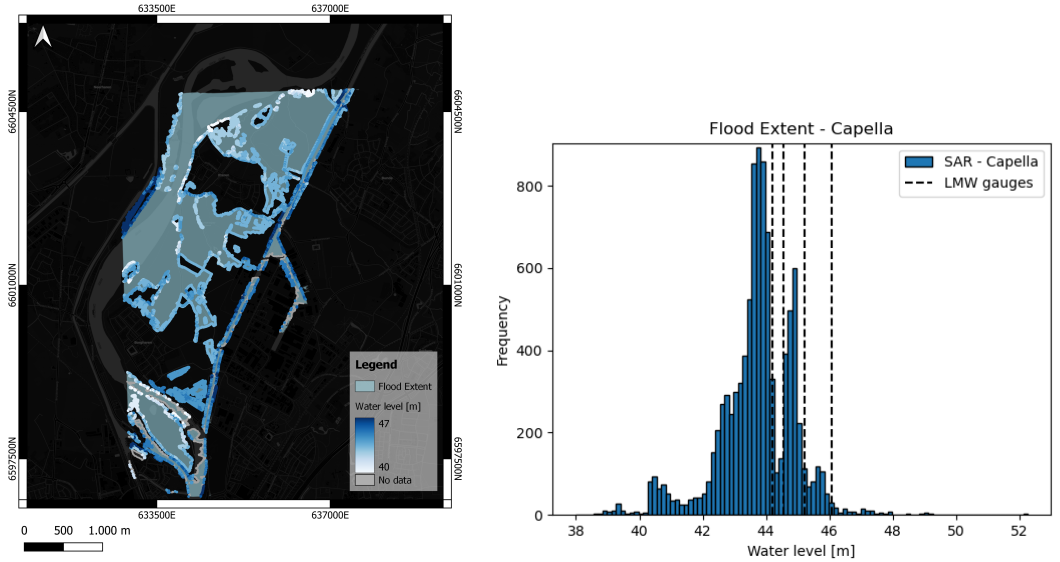
The error matrices of the flood extents indicate the imbalance in the dataset, this was also seen in the error matrix of between the Capella Space and modeled datasets. When looking at performance parameters of Table 4.4 and Table 4.5, the precision increases between the two results but the recall decreases. The precision is higher in the first matrix due to overestimation of the SAR-based extent that can also be seen in Figure 4.13. Comparing the SAR-based extent to the modeled results at 07:00 UTC, it can be seen that more flooded pixels of the SAR are correctly classified by the model as well. The recall decreases as the modeled results increase in flood extent, that is not fully captured by the SAR results. The specificity is high in both cases, indicating that the amount of non-flooded pixels of the modeled results are being consistently detected. The F1-scores are similar, but lower compared to the F1-score of the error matrix between Capella Space data and the model.

The error matrix and performance parameters of the combined flood extent show a high precision and specificity, indicating that the flooded pixels of the SAR-based extent are consistent with the modeled results. The recall is lower than previous results, which can also be seen in Figure 4.14b. Based on the results, the SAR-based flood extent shows underestimation with respect to the modeled results. The lower values are also translated into the F1-score of 0.62, which is the lowest value of all computations.

4.3. Capella Space - Water Level estimation

As explained in chapter 3, when computing the flood extent, it is important to look into the outer edges of the polygon. Therefore, the flood extent polygons are sampled every 10 m along the outer edges.

The results of the flood extent derived from Capella Space data can be computed by taking the extracted DTM heights as representative values for the water level estimates. The results are shown in Figure 4.18 with an histogram of all computed water levels within the area, together with the corresponding measured water levels at the LMW gauges shown in Figure 2.17a.



(a) Sampled results displayed on the map.

(b) Histogram of all the results.

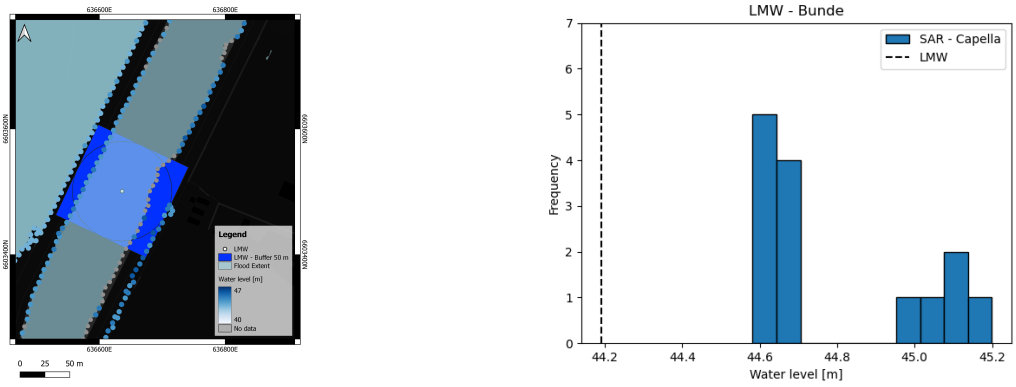
Figure 4.18: Capella Space water level estimates computed every 10 m sampled along the edges of the flood extent polygons.

4.3.1. LMW validation

Several LMW gauges are located in the AOI, as shown in Figure 2.17a. These gauges contain measurements of the water levels every 10 minutes during the flood event. As these gauges are a valuable measurement tool for Rijkswaterstaat, it is of importance to compare the measured values of the gauges to the SAR-based water levels. The accuracy required by Rijkswater is 20 cm.

The gauges intersecting the flood extent were taken into account, the other locations were either too far away from the river or no flood extent was captured by the SAR data. Beneath, all LMW gauges are shown on the map with the corresponding SAR-based water level results and a frequency distribution of all the representable estimates, the statistics are listed in Table 4.7. Differences in water levels between the sources are discussed together with the water levels derived from Sentinel-1.

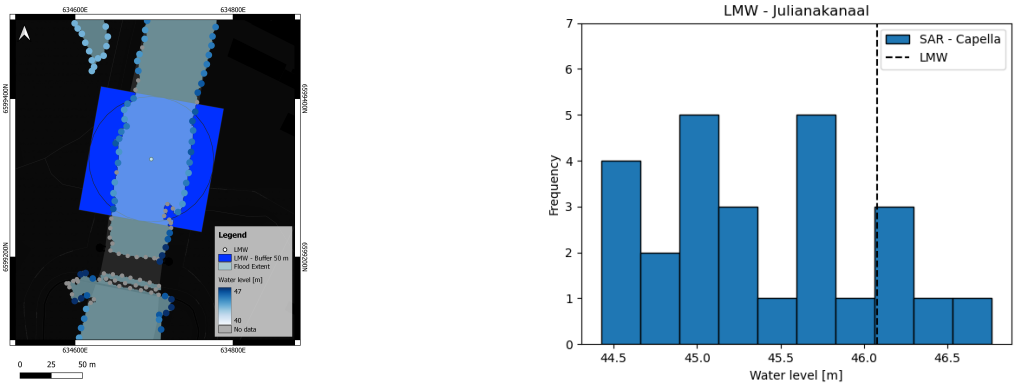
For the results computed at the location of Lanaken, the two highest bins in the distribution were used to generate the average and median values of the water levels. The area clearly shows a lot of variation where the two upper bins can be considered to be related to the maximum water level.



(a) LMW Location.

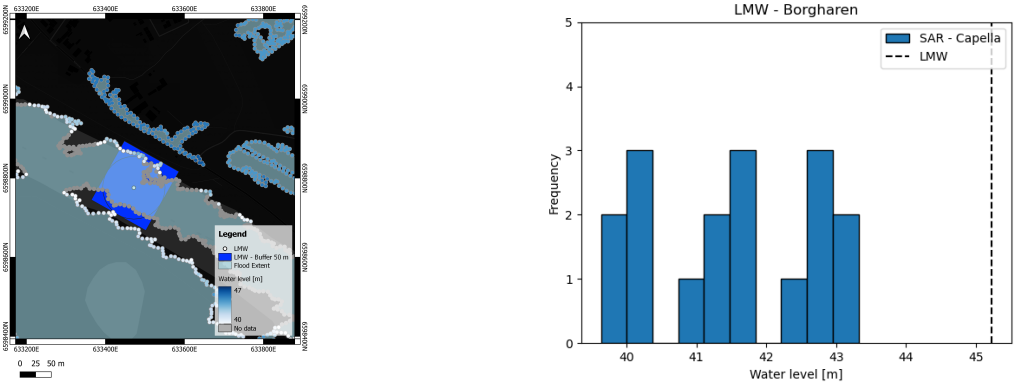
(b) Histogram of the SAR-results.

Figure 4.19: Location of the LMW gauge at Bunde with the corresponding flood extent and water level estimates derived from the Capella Space data.



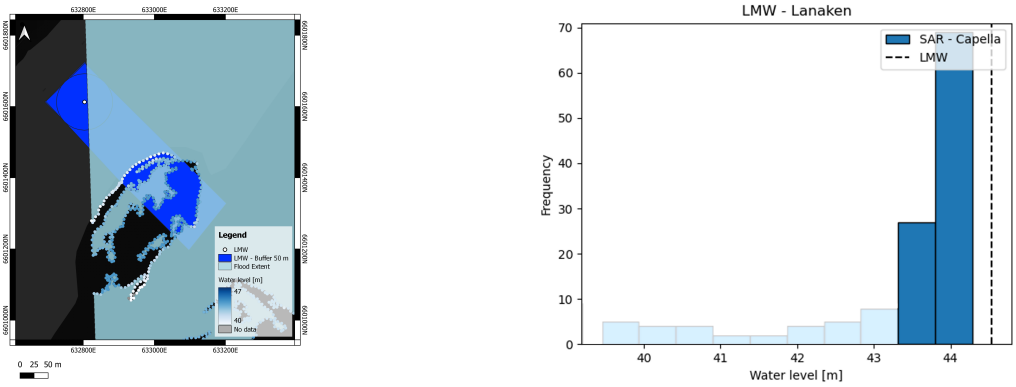
(a) LMW Location. **(b)** Histogram of the SAR-results.

Figure 4.20: Location of the LMW gauge at Borgharen Julianakanaal with the corresponding flood extent and water level estimates derived from the Capella Space data.



(a) LMW Location. **(b)** Histogram of the SAR-results.

Figure 4.21: Location of the LMW gauge at Borgharen Dorp with the corresponding flood extent and water level estimates derived from the Capella Space data.



(a) LMW location. **(b)** Histogram of the SAR-results.

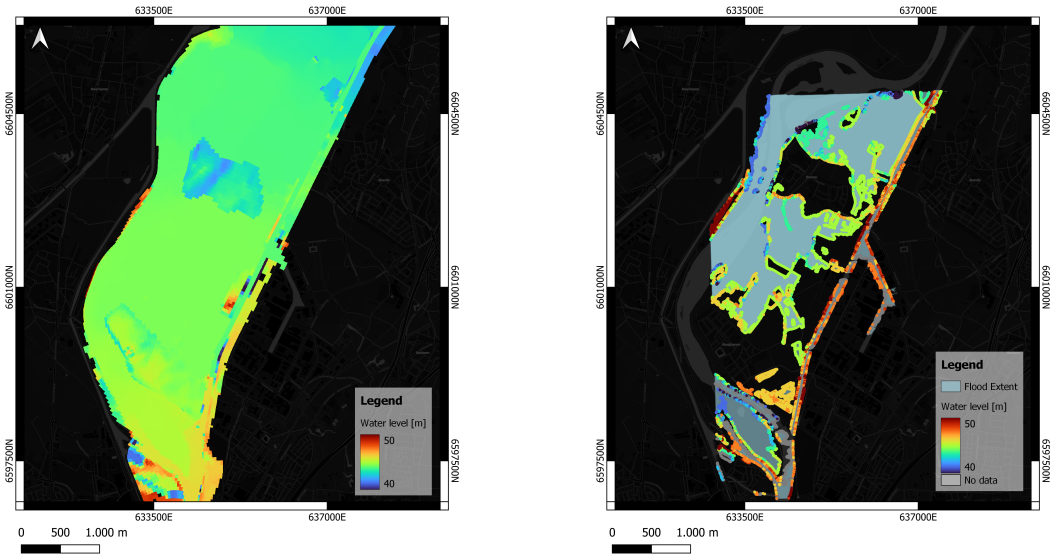
Figure 4.22: Location of the LMW gauge at Lanaken with the corresponding flood extent and water level estimates derived from the Capella Space data.

Table 4.7: Comparison between LMW water level measurements and the Capella Space derived water level estimates at 23:30 UTC. All water level measurements are provided in meters.

Location	LMW	Avg. SAR	Median SAR	Std. Dev.	Avg. Slope [^{o}]
Bunde	44.19	44.93	44.66	0.221	5.24
Julianakanaal	46.08	45.37	45.28	0.373	6.88
Borgharen Dorp	45.23	41.61	41.80	0.646	13.59
Lanaken	44.54	44.07	44.09	0.151	1.70

4.3.2. Model validation

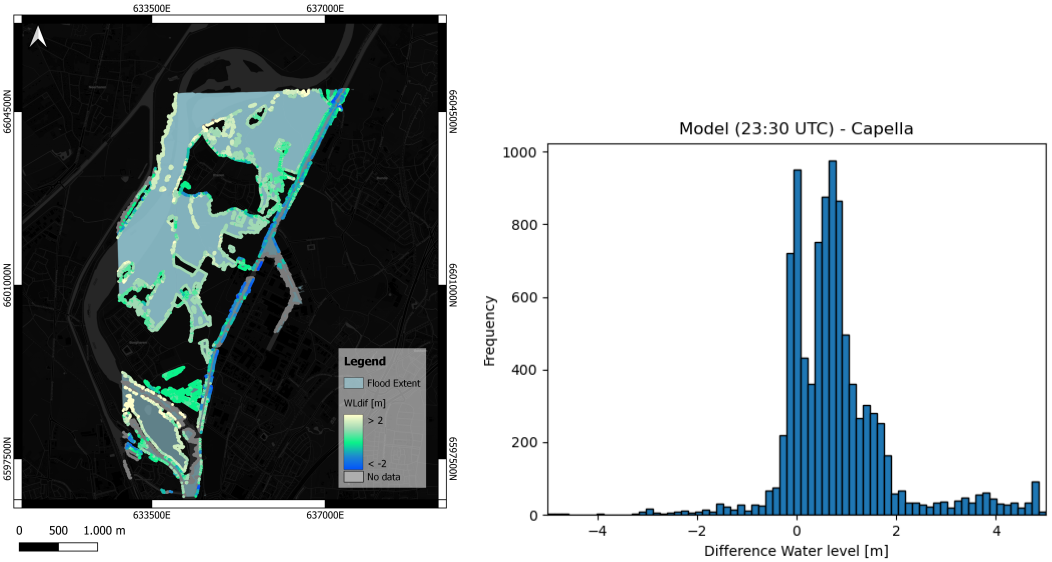
The hydrodynamic model of Deltares simulates the flood event with different parameters, among others the water depth and the water level. The estimated water levels derived from the flood extent were compared to the results of the model, at the same time of the satellite acquisition, which is at 15/07/2021 23:30 UTC. Figure 4.23 shows the water level data from both sources with the same colorscale and Figure 4.24 shows the differences in water level based on Equation 3.12 with the corresponding histogram. Negative values indicate higher SAR-based water levels and positive values indicate higher modeled values. The histograms shows a peak between 1-2 m difference, indicating underestimation of SAR-based results compared to the model.



(a) Simulated by the Deltares model.

(b) SAR-based water levels from Capella Space data.

Figure 4.23: The computed water levels within the AOI.



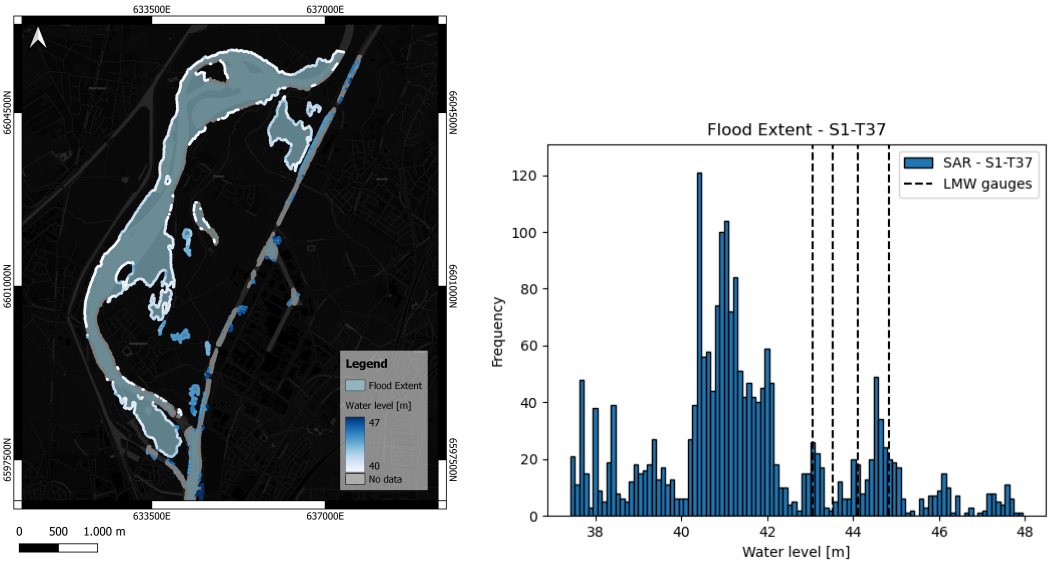
(a) AOI.

(b) Histogram of the results.

Figure 4.24: Differences between the water level results from the hydrodynamic model of Deltares and the SAR-based water levels from the Capella Space data.

4.4. Sentinel-1 - Water Level estimation

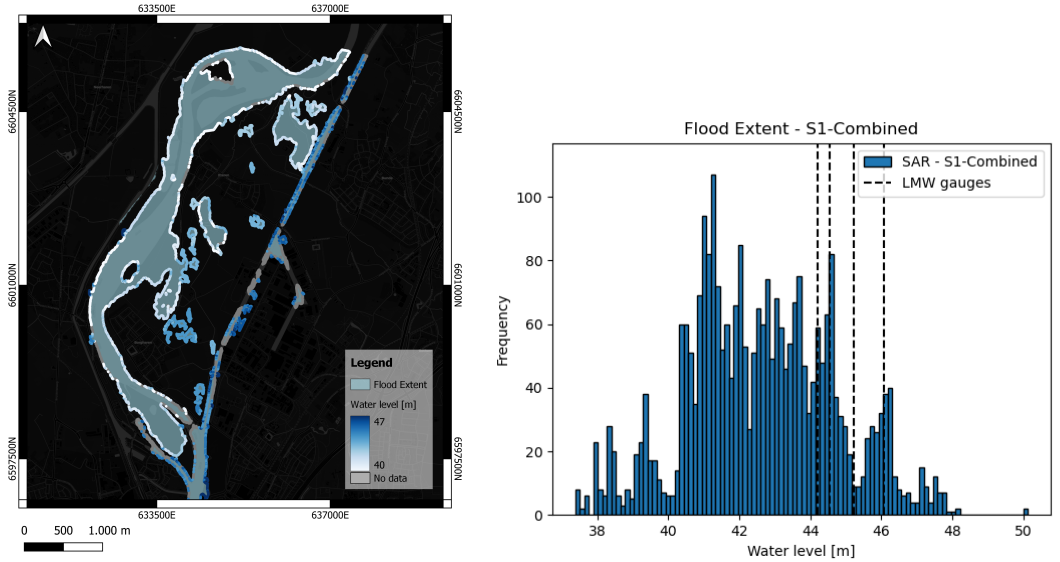
The estimated water levels from the Sentinel-1 extent are derived in a similar way as applied to the Capella Space data. The Sentinel-1 water level results can be computed using the extracted DTM heights, which represent the water level estimates. The results are shown in Figure 4.25 for the T37 stack and Figure 4.26 for the Combined flood extent, with an histogram of the derived water levels within the AOI.



(a) Sampled results displayed on the map.

(b) Histogram of all the results.

Figure 4.25: Water levels estimates from Sentinel-1 data on 15/07/2021, sampled every 10 m along the edges of the flood extent polygons.



(a) Sampled results displayed on the map.

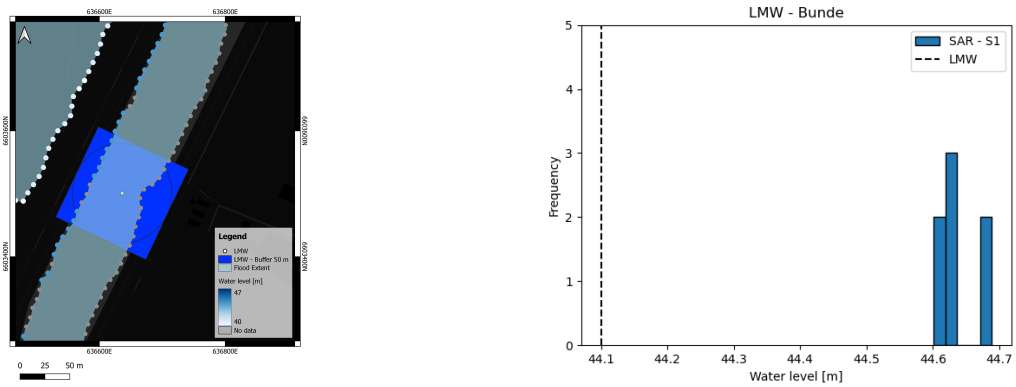
(b) Histogram of all the results.

Figure 4.26: Water levels estimates from Sentinel-1 data of the combined acquisitions on 15/07/2021 and 17/07/2021, sampled every 10 m along the edges of the flood extent polygons.

4.4.1. LMW validation

The comparison with LMW gauges located in the AOI (see Figure 2.17a), have been performed with the water level estimates acquired on 15^{th} of July. The use of SAR with respect to LMW water level measurements is mainly of interest in the timing component. For this validation it is important to evaluate the agreement between the SAR-based water levels and levels measured at the LMW gauges at the time of acquisition of the satellite. In this case, it is unnecessary to validate the results to the combined flood extent with two different acquisition dates. The gauges where outliers were detected on the representative side of the canal or river are taken into account. If the LMW locations are not mentioned in the results, there were no outliers detected around the location. Beneath, the LMW gauges are shown with the corresponding results of the Sentinel-1 derived flood extent and water level estimates. The representable estimates are also listed in Table 4.8.

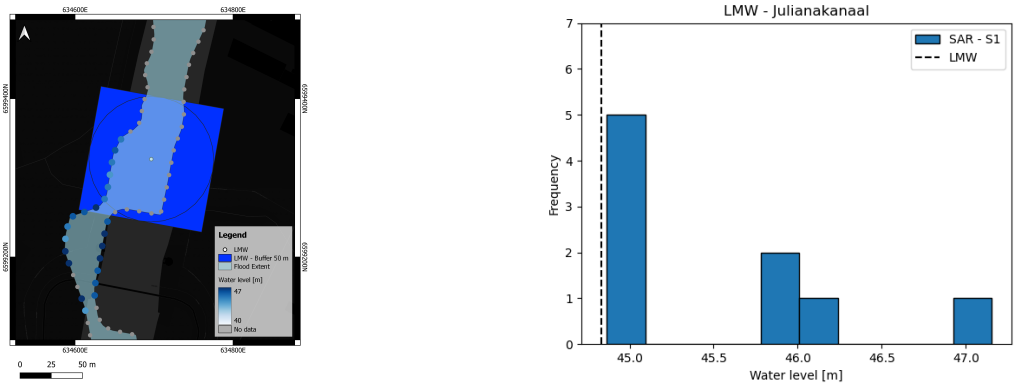
The statistics are computed similar as mentioned for the Capella Space data. For the results regarding the location of Lanaken, the three upper bins from the distribution have been used to generate the average and median values of the water levels. The area clearly shows a lot of variation where the three upper bins can be considered to be related to the maximum water level.



(a) LMW Location.

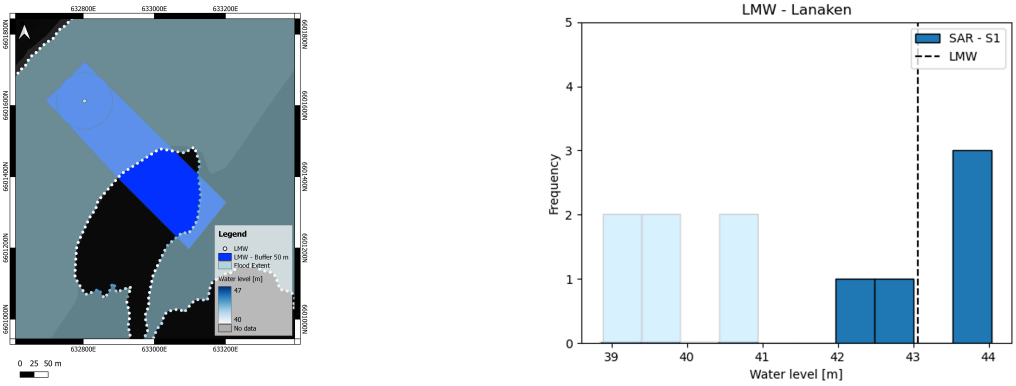
(b) Histogram of the SAR-results.

Figure 4.27: Location of the LMW gauge at Bunde with the corresponding flood extent and water level estimates derived from Sentinel-1 data.



(a) LMW Location. **(b)** Histogram of the SAR-results.

Figure 4.28: Location of the LMW gauge at Borgharen Julianakanaal with the corresponding flood extent and water level estimates derived from Sentinel-1 data.



(a) LMW location. **(b)** Histogram of the SAR-results.

Figure 4.29: Location of the LMW gauge at Lanaken with the corresponding flood extent and water level estimates derived from Sentinel-1 data.

Table 4.8: Comparison between water level measurements from LMW and the Sentinel-1 derived water level estimates at 05:50 UTC. All water level measurements are provided in meters.

Location	LMW	Avg. SAR	Median SAR	Std. Dev.	Avg. Slope [^{o}]
Bunde	44.10	44.64	44.63	0.04	4.00
Julianakanaal	44.83	45.78	45.58	0.585	10.21
Lanaken	43.06	43.05	43.32	0.393	12.22

The location of the LMW gauges at Bunde is shown in Figure 5.6a. The gauge showed an increase of 0.09 m in water level between the two satellite acquisitions on 15/07/2021. The slope is relatively low for both flood extents, and accordingly the standard deviations are 22.1 cm and 4 cm for Capella Space and Sentinel-1 data respectively.

The location of the LMW gauge at Borgharen Julianakanaal is shown in Figure 5.6b. The gauge showed an increase of 1.25 m in water level between the two satellite acquisitions on 15/07/2021. The slopes are a steeper, which also translated into the higher standard deviations of 37.3 cm and 58.5 cm for

Capella Space and Sentinel-1 data respectively.

The location of the LMW gauge at Borgharen Dorp is shown in Figure 5.6c. The gauge showed an increase of 1.71 m in water level between the two satellite acquisitions on 15/07/2021. The slopes around this location are among the steepest in the area as the river is surrounded by dikes. Sentinel-1 ATSA analysis did not detect any significant outliers in the area. The extent derived by Capella Space shows underestimation of the water level and with the highest standard deviation of 64.6 cm.

The location of the LMW gauge at Lanaken is shown in Figure 5.6d. The gauge showed an increase of 1.48m in water level between the two satellite acquisitions on 15/07/2021. The location contains a small hill around which the flood extents are detected. The differences in slope are large between the water levels from Sentinel-1 and Capella Space. It is expected that the extent of Capella Space is located on top of the hill, explaining the smaller slope value. The difference in standard deviation are 15.1 cm for Capella Space and 39.3 cm for Sentinel-1.

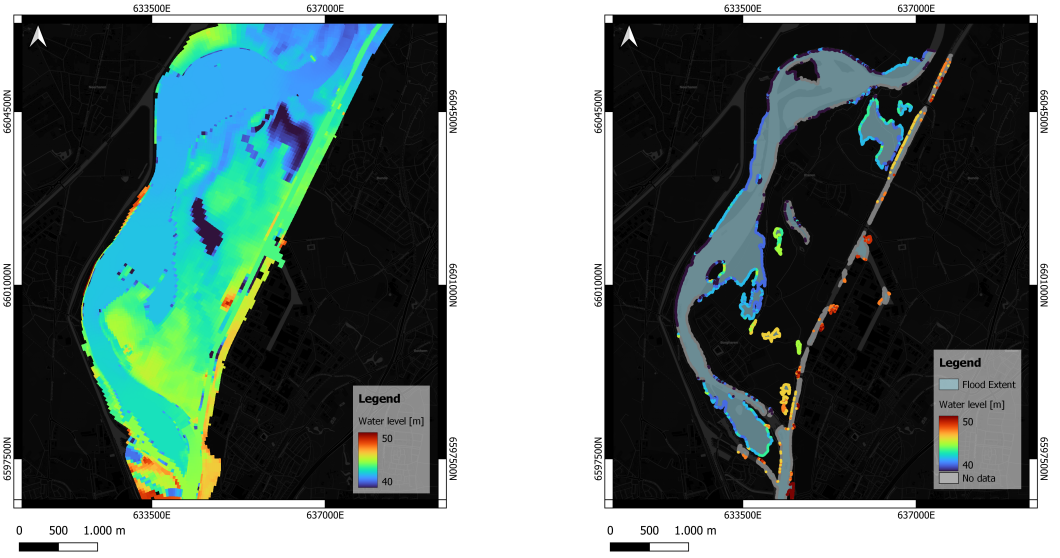
The differences between the LMW gauge and the SAR estimates can be deduced from Table 4.7 and Table 4.8, and are summarized in Table 4.9. The differences are further discussed in the next chapter.

Table 4.9: Comparison between LMW water level measurements and SAR-based water levels. All water level measurements and standard deviation are provided in meters.

Location + SAR	LMW - Avg.SAR	LMW - Median SAR	Std. Dev. SAR	Avg. Slope[^{o}]
Bunde				
Sentinel-1	-0.54	-0.53	0.040	4.00
Capella	-0.74	-0.47	0.221	5.24
Julianakanaal				
Sentinel-1	-0.95	-0.75	0.585	10.21
Capella	0.71	0.80	0.373	6.88
Borgharen				
Sentinel-1	No data	-	-	-
Capella	3.62	3.43	0.646	13.59
Lanaken				
Sentinel-1	0.01	0.26	0.393	12.22
Capella	0.43	0.45	0.151	1.70

4.4.2. Model validation

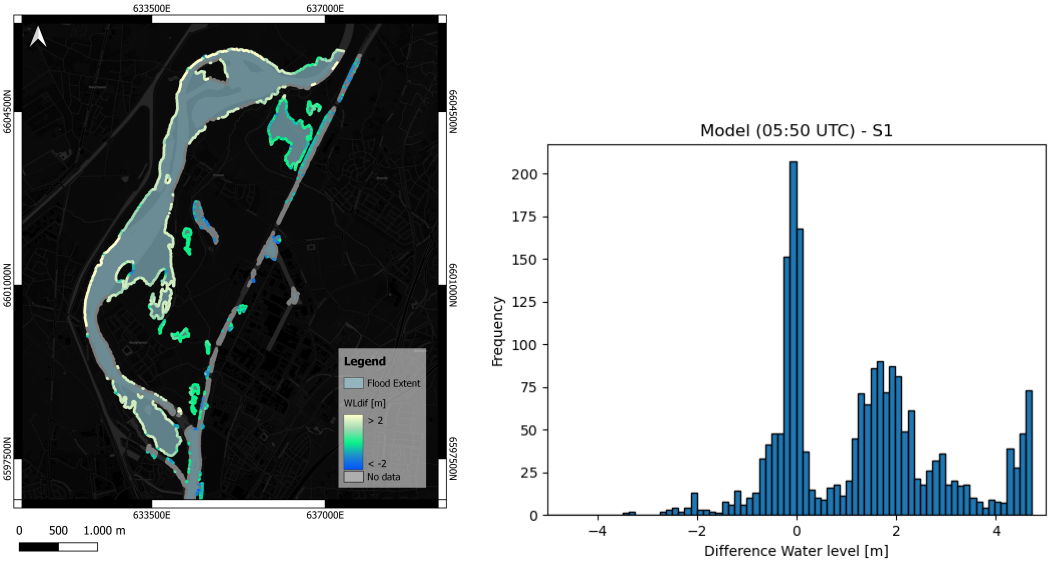
As mentioned before, the hydrodynamic model of Deltares simulates the flood event with different parameters, among others the water depth and water level. Comparing the flood extent of 15^{th} of July to the same time stamp in the model can provide insights into the agreements between SAR data with the modeled data. Figure 4.30 and Figure 4.32 show the water level data of both sources at the time of the satellite acquisition at 15/07/2021 05:50 UTC and at 15/07/2021 07:00 UTC, respectively. Figure 4.31 and Figure 4.33 visualize the differences between modeled results and the SAR-based water levels with Equation 3.12. In case WL_{dif} is negative, this indicates there is an overestimation of the SAR-based water levels. Vice versa, if WL_{dif} is positive, it represents an underestimation of the SAR-based water levels with respect to the modeled water level results. The histograms of Figure 4.31 and Figure 4.33 show a peak around 0 m and 2 m difference, this indicates that the area shows some alignment between the two datasets and underestimation by the SAR-based results. By looking at the map, most of the underestimated areas are located around the edges of the river where high vegetation is present.



(a) Simulated by the Deltares model.

(b) SAR-based water levels from Sentinel-1.

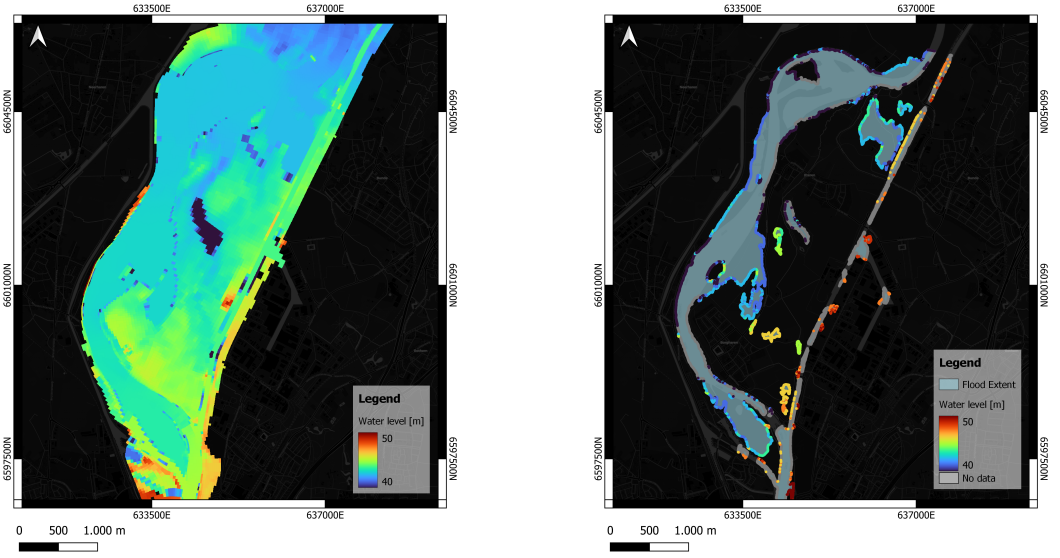
Figure 4.30: Water levels within the AOI on 15/07/2021 at 05:50 UTC.



(a) AOI.

(b) Histogram of the results.

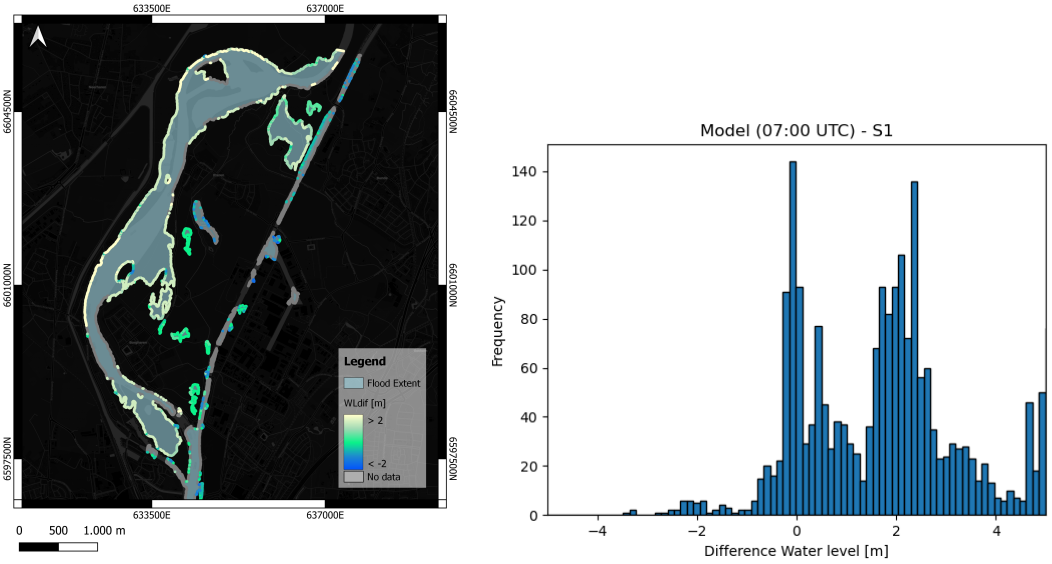
Figure 4.31: Differences between the water level results from the hydrodynamic model of Deltares and the SAR-based water levels from the combined Sentinel-1 acquisitions at 15/07/2021 05:50 UTC.



(a) Simulated model results at 07:00 UTC.

(b) SAR-based water levels from Sentinel-1.

Figure 4.32: Water levels within the AOI on 15/07/2021, comparing the Deltares model to the SAR-based water levels.

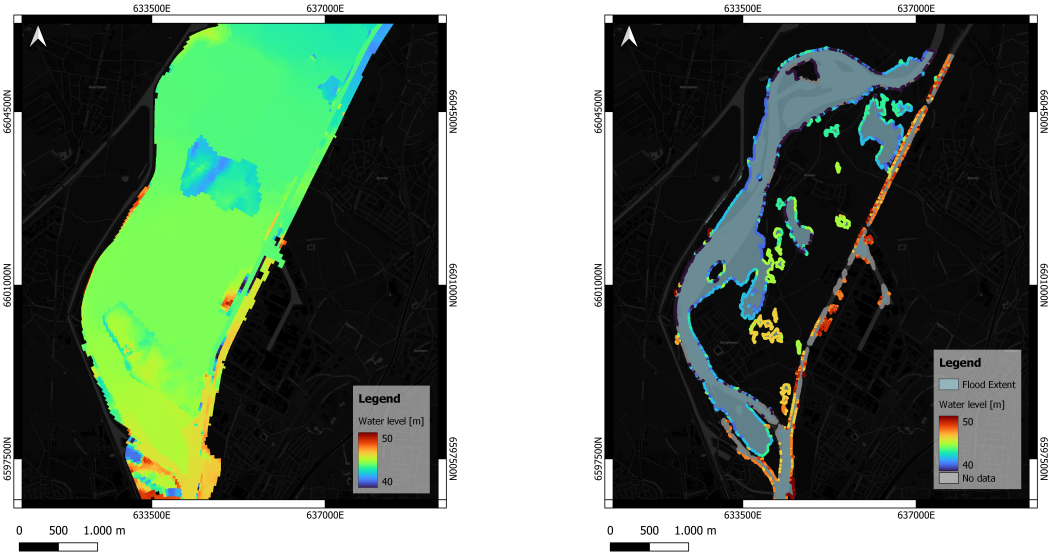


(a) AOI.

(b) Histogram of the results.

Figure 4.33: Differences between the water level results from the hydrodynamic model of Deltares at 15/07/2021 07:00 UTC and the SAR-based water levels from the combined Sentinel-1 acquisitions 15/07/2021 05:50 UTC.

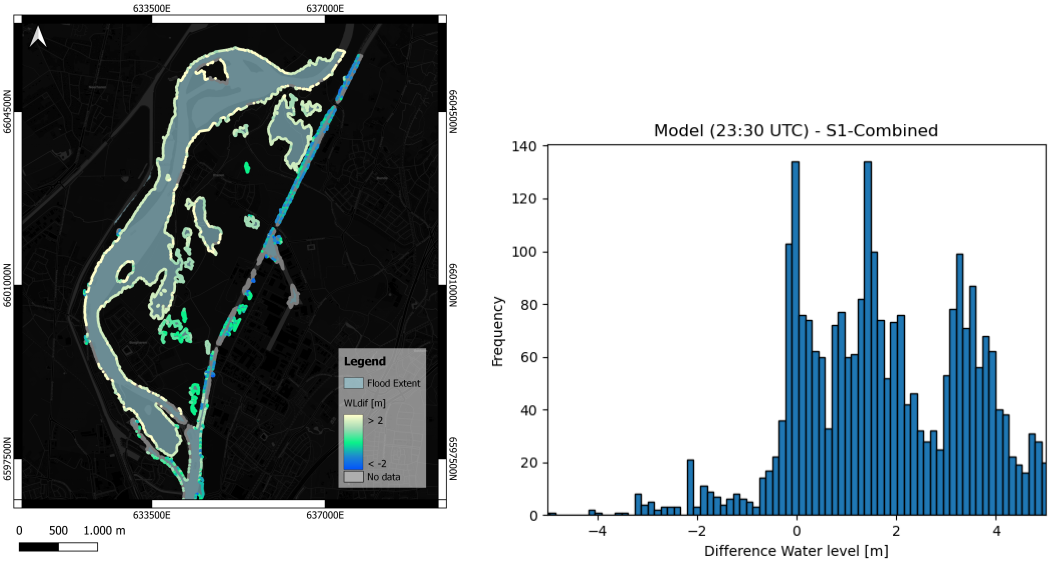
The combined flood inundation extent of the detections from the 15^{th} and 17^{th} of July are also compared to the peak of the flood from the model results. The peak is estimated at approximately 23:30 UTC. A comparison is done to validate whether changes in soil moisture or affected land can be detected two days after the peak. Combining the extent of both days can potentially provide the extent of the entire area that was affected by the floods. The derived water levels and the differences between the model data is computed equivalently to the other approach and the results are displayed in Figure 4.34 and Figure 4.35.



(a) Simulated by the Deltares model.

(b) SAR-based water levels from Sentinel-1.

Figure 4.34: Water levels within the AOI at the time of the peak of the flood: 15/07/2021 23:30 UTC.



(a) AOI.

(b) Histogram of the results.

Figure 4.35: Differences between the water level results from the hydrodynamic model of Deltares at 15/07/2021 23:30 UTC and the SAR-based water levels from the combined Sentinel-1 acquisitions.

4.5. Summary

All results have been displayed in this chapter. First, the computation of the flood extents that have been compared to the modeled results at the representative time stamps. Thereafter, the results of the water level estimations have been shown. The results of the SAR-based water levels are compared to the LMW gauge measurements and to the modeled results at the corresponding timestamps of the Deltares model. On average, the standard deviation of the water level points is 0.141 m for Capella Space, 0.156 m for Sentinel-1 extent of the T37 stack, and 0.210 m for the combined Sentinel-1 extent of T37 and T161 stack. Visual validation of the aerial images and LiDAR derived Water-Land boundary lines will be discussed in the next chapter, along with the examination of all other results.

5 Discussion

In chapter 4 the results of the SAR-based flood extents and water levels have been presented for two selected SAR data sources. The results are presented with a brief explanation, along with a validation against the LMW gauges and Deltares model data. In this chapter, the results will be discussed in more detail and other validation data will be further assessed. Finally, the advantages and shortcomings of SAR data and real-life applications of the methods are discussed.

5.1. Flood extent

An accurate estimate of the flood extent is an important and crucial step in estimating water levels. For both SAR data sources, the extents have been computed and displayed in chapter 4. In this section, the accuracy, performance and inefficiencies are discussed. The flood extents are also compared to aerial images and LiDAR data derived during the flood. The performance parameters of the error matrices are displayed in Table 5.1.

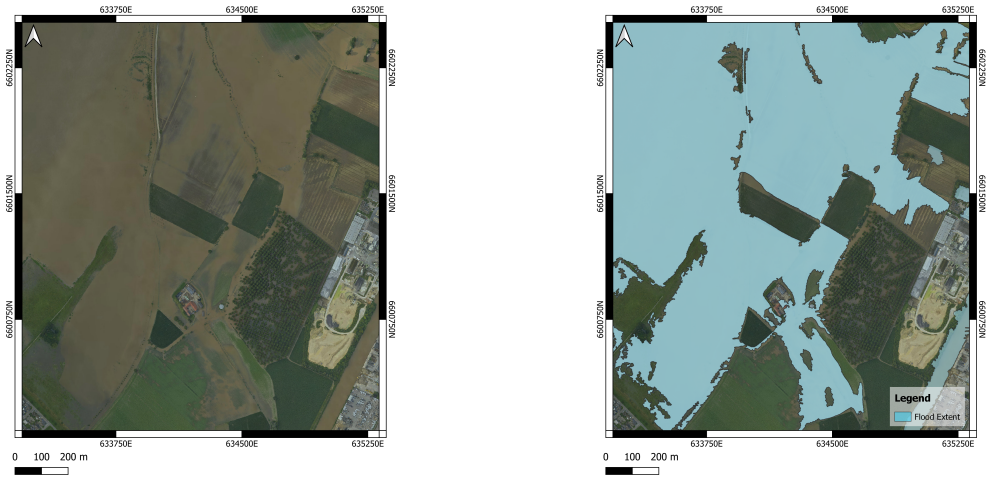
Table 5.1: Performance parameters of all error matrices between modeled results and SAR-based flood extents.

SAR	Accuracy	Precision	Recall	Specificity	F1-Score
Capella Space	0.73	0.92	0.67	0.87	0.78
Sentinel-1 T37 (05:50 UTC Results)	0.94	0.82	0.68	0.98	0.74
Sentinel-1 T37 (07:00 UTC Results)	0.93	0.87	0.64	0.98	0.74
Sentinel-1 T37 + T161	0.86	0.97	0.45	0.995	0.62

5.1.1. Capella Space

The flood extent from Capella Space is computed using the thresholding method, where the threshold was determined by backscatter profile analysis. In the sampled regions of flooded areas with low vegetation, it was observed that the flood water ranges overlap with permanent water. This confirms that floods can be detected using specular reflection phenomena. Flood water can become more complicated to detect through debris, emergent vegetation or scattering in the waves of the water. These variables broaden the range of backscatter values, making classification more complex. Locations with higher vegetation, such as corn fields, generate more double bounces when flood water is present. These fields had an opposite spectral signature with higher backscatter values compared to the areas where more specular reflection was found.

The interaction between SAR signal and vegetation also influences the classification of water. Capella Space data is X-band data, indicating that the signal has a limited penetration depth into the vegetation. The interactions between the signal and the tree leaves can produce artifacts that reduce the accuracy of the flood extent. Figure 5.1 shows the effects of vegetation on the underestimation of the flood extent.



(a) Aerial image.

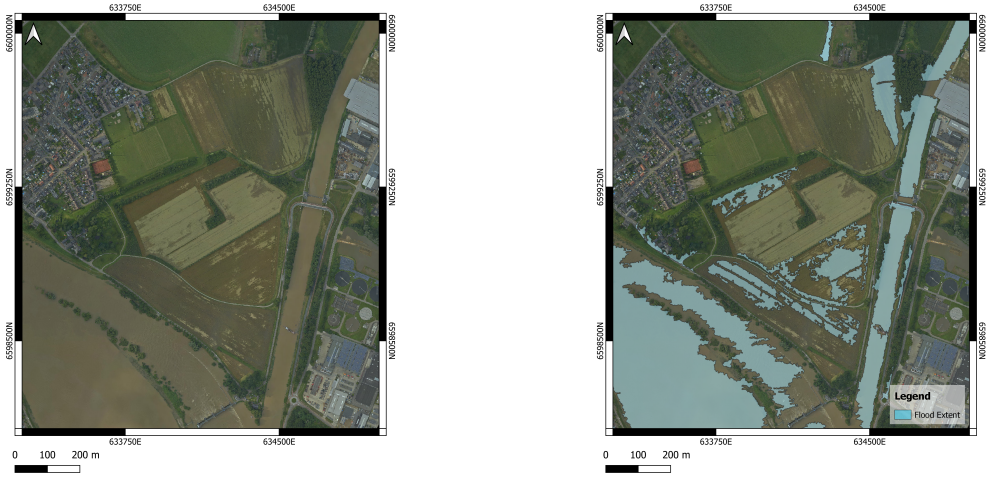
(b) Aerial image and flood extent.

Figure 5.1: Example of influences of high vegetation on the classification of flooded pixels.

The threshold was applied to a single image, this method is prone to false positives. Locations with highways, airport runways, shadows of large trees or buildings were falsely classified as flooded. These structures within a SAR image have smooth surfaces with backscatter characteristics similar to those of water bodies. This similarity complicates the classification using a single image. An example of false positive classification of the airport runway is shown in Figure 4.9. Using a baseline image, the number of false positives can be reduced for stationary areas.

Using the post-event image, an example was shown to evaluate the effects of having a baseline image. The baseline image was able to reduce false detections, however, it may be difficult to implement in real-life applications. The latest high-resolution satellites often deliver data on-demand, within a few hours. This is beneficial for quick image acquisition, but the disadvantage is that there is limited historical data available. Another downside is that the images have to be taken over the same area, the co- and post-event image for this case study overlapped for 25-50% of the images (see Figure 4.7). Concerning real-life applications, it is more probable that a single image is available over an AOI.

From Figure 4.6 can be deduced that 67% of the flood extent computed by the Deltares model was correctly classified by the SAR extent, as well. The F1-score is 0.78 which is the highest score considering all computations. The precision of the flood extent is 0.92, meaning that most of the SAR-based flood extent is within the extent of the model. The specificity is 0.87 which indicates that there are locations outside the extent that have been classified as flooded by the SAR. Overestimations from SAR are mostly related to puddles in the fields, this effect can be seen in Figure 5.2. The recall is 0.67, indicating underestimation compared to the model. Underestimated areas are mostly related to densely vegetated areas or the edges of the flood extent. Underestimation on the edges of the extent could also be related to difference in resolution cells between model (20 m x 40 m) and SAR (0.7 m x 0.7 m), as displayed in Figure 5.3. The flood extent derived from Capella Space contains more details compared to the modeled results which could lead to less classifications as the mode interpolates the data over larger grid cells.



(a) Aerial image.

(b) Aerial image and flood extent.

Figure 5.2: Example of puddle formation resulting in classification of flooded pixels.

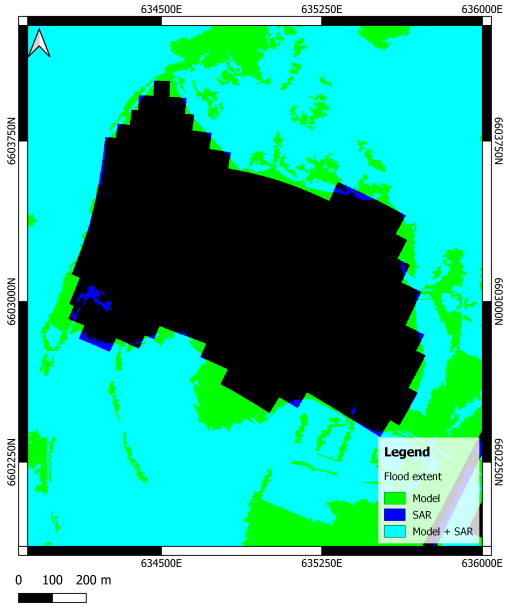
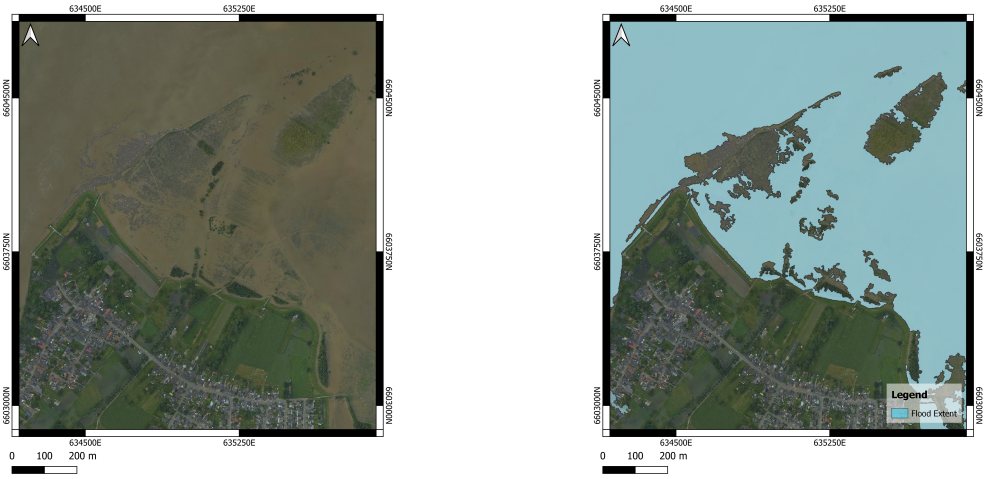


Figure 5.3: Zoom-in on the edges of the model displaying the differences in resolution.

Aerial photographs have also been taken over the area and captured the AOI just after the peak of the flood on 16/07/2021 between 16:15 and 17:29 UTC. In Figure 5.14, three different parts of the AOI display the differences and consistent areas of the flood extent. The SAR-based flood extent is in most areas in agreement with the aerial images. Some missing areas are within fields that contain high vegetation, such as corn and high trees. In the southern part of the AOI, increased volume scattering due to high vegetation is clearly visible. Another area that is less accurately detected corresponds to a region where debris is visible in the aerial image. Debris is known to affect the radar signal, increase scattering, and reduce the specular reflection effect, an example is given in Figure 5.4.

Detections within the SAR image that show puddle formation in the fields are considered to be falsely classified as flooded. The detection of water is not false, however, this type of water damage would not be considered to be part of the flood extent. The distinction between puddles on the streets or fields and the flood extent, is one of the complexities of comparing this technique to other data types. The

model data may not simulate puddle formation, and river gauges do not measure on the land, but these effects will show in the SAR image.



(a) Aerial image.

(b) Aerial image and flood extent.

Figure 5.4: Example of the influences of debris in classification of flooded pixels.

5.1.2. Sentinel-1

The flood extents of Sentinel-1 were derived from outliers detected by the ATSA approach. This approach focuses on using all historical data to detect changes of each specific date, per pixel. The SLC data of the T37 and T161 stacks were used to create two flood extents. Using the ATSA method, the influences of roads or shadows become less apparent as no temporal changes in backscatter arise (assuming that the areas are not flooded). This method is prone to false detections by dynamic locations such as industrial terrains that are consistently varying over time. It was expected that detection of flooded pixels would be mostly related to negative outliers related to specular reflection. Especially in open and flat agricultural areas, water acting as a specular reflection shows as a negative spike in the time series.

Detections become complex for shallow water bodies such as the canals. A limited number of detections were captured along the canals, which can be explained by the size of the canals (50 m) and the spatial resolution of the Sentinel-1 satellite. The application of a speckle filter is an advantage considering the influence of speckle noise. However, due to the application of a despeckling filter the spatial resolution of the satellite is even further decreased.

Negative outliers are related to the characteristics of water acting as a specular reflection. Positive outliers could provide additional information on the total inundated area. When water reaches a certain location and subsequently recedes over time, the soil characteristics are expected to change for several days. Higher soil moisture or the presence of debris can show a distinct reflection compared to normal conditions. These effects can lead to an increase in backscatter, due to an increase in volume or double bounce scattering effect of the SAR signal. By expecting that these changes are significant enough to detect, the flood inundation extent is composed of the two Sentinel-1 acquisitions (15^{th} and 17^{th} of July). The effects mentioned above are mainly related to low surface water levels and return to normal conditions after the water recedes into the ground. Some locations of submerged areas appear on the map from the 17^{th} July as positive outlier detections, confirming the theory of increased backscatter in vegetated areas with low surface water.

Figure 4.15 shows that 68% of the flood extent computed by the Deltares model (at 05:50 UTC) was correctly classified by the SAR extent. The F1-score is 0.74, which is lower than the F1-score of the flood extent from Capella Space data. The precision is 0.82, meaning that the SAR-based flood extent showed more overestimation compared to the model as Capella Space. The specificity remains high in this dataset, this is because of the imbalance in the dataset that the amount of False Positives is not significant enough to change the score. The lower value of precision is expected as Figure 4.13

displays an underestimation of the modeled results due to inconsistent timing. The recall is 0.68, also indicating an underestimation of the SAR results compared to the model. The underestimation is mainly related to the edges of the modeled extent, that are located at the edges of the river that are covered with high trees, two locations are displayed in Figure 5.5. Also the spatial resolution of Sentinel-1 is higher compared to the model grid, similar as displayed in Figure 5.3.

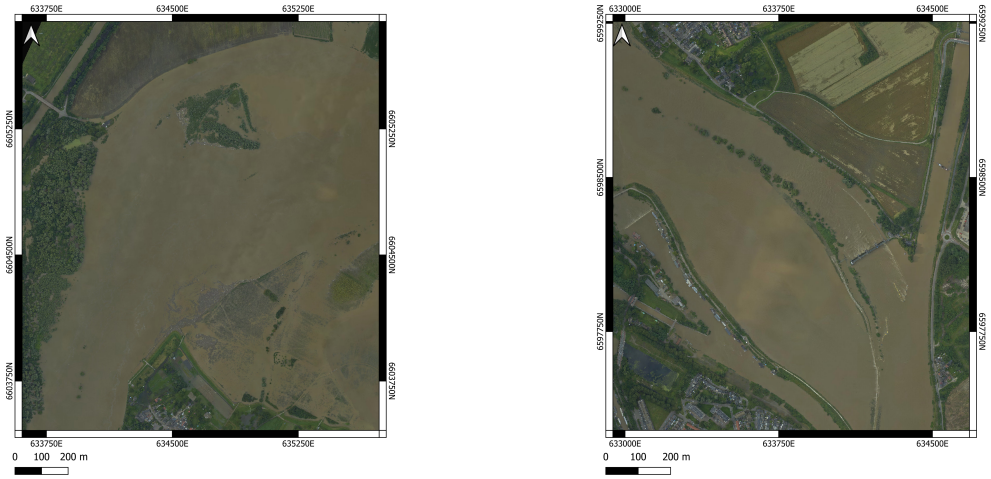


Figure 5.5: Example of the trees located on the edges of the Meuse river.

The results changed to 64% correct classifications compared to the modeled results at 07:00 UTC. The F1-score is similar to the first with 0.74, however, the precision increased to 0.87 and the recall decreased to 0.64. The precision increased as the SAR-based flood extent showed less overestimation compared to the later modeled results. The recall decreased due to extra underestimation at the edges of the model, potentially caused by some densely vegetated areas with high trees or presence of debris which resulted in false negatives.

The most significant differences between the the results at 05:50 UTC time and 07:00 UTC time, can be seen in Figure 4.13. There are several areas that show the increased flood extent between the moments, where the results of 07:00 UTC show a higher agreement to the SAR-based flood extent. This indicates that the modeled results are approximately 1 hour and 10 minutes behind on timing of the inflow of water.

The combined flood inundation map showed 45% correctly classified pixels of the modeled results at 23:30 UTC (peak of the flood). The F1-score is 0.62, the lowest value of all computations. The precision is 0.97 and the specificity 0.995, both indicate that the number of False Positives is limited. The classifications of flooded pixels and non-flooded pixels show good agreement with the modeled results, indicating that the computed flood extent is actually flooded. The recall 0.45, this suggests that the SAR-based flood extent shows underestimation compared to the model. This can also be seen in Figure 4.14b, where an extensive part of the model extent is not covered by the extent of the SAR-based results. The higher number of False Negatives compared to the results from Capella Space data is mainly due to the timing of the SAR acquisition that does not capture the peak of the flood.

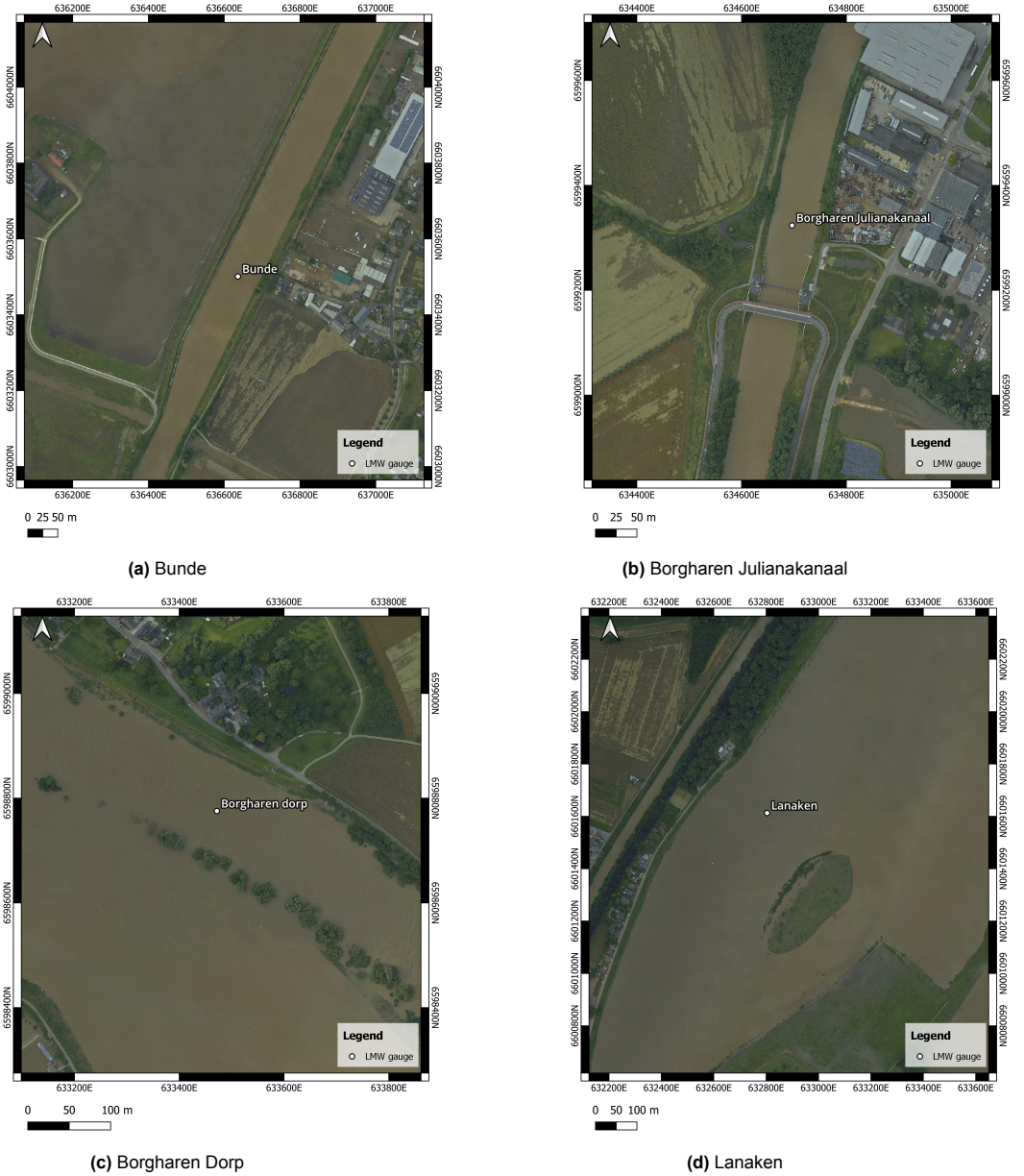
Aerial images were taken on the day in between the two Sentinel-1 acquisitions. In Figure 5.14, three areas are shown to compare the Sentinel-1 derived flood extents. The extent of the T37 stack on 15^{th} of July is based on the situation right before the peak of the flood. The areas are enclosed within the flooded area that can be seen in the aerial images. The water levels were lower at the time of SAR acquisition, therefore an adequate comparison cannot be made between the images. The combined SAR extent shows an increased consistency to the flooded area. The additional flooded locations are mostly along the canals and fields, but do not approximate the maximum extent seen in the aerial images similar as to the modeled extent. The locations that detected additional inundation extents are located inside the flooded areas seen in the aerial images. This indicates that the additional inundated fields detected are True Positives and are related to the presence of flood water.

5.2. Water level estimates

The estimated water levels from Capella Space and Sentinel-1 data are shown in Figure 4.18, Figure 4.25 and Figure 4.26. The histogram of the results indicate whether the measured water levels fall within the range of heights measured by the LMW gauges. Heights that fall outside these ranges can indicate either over- or underestimation of the flood extent. Over- or underestimations can be interpreted as water reaching outside the banks of the rivers or filling of the floodplains. As opposed to the gauges, the SAR-based water levels are an overview of the total AOI while the gauges refer to specific locations. For this reason, direct conclusions cannot be drawn from the histogram. The results can provide an initial approximation of the distribution of water levels within the AOI.

As the SAR-based extents are acquired at specific moments in time, the extracted water levels were compared to the LMW measurements at the corresponding time.

5.2.1. LMW validation



(a) Bunde

(b) Borgharen Julianakanaal

(c) Borgharen Dorp

(d) Lanaken

Figure 5.6: Aerial images during the flood of the LMW locations and their locations projected to the center line of the Meuse river.

The differences between the LMW and SAR estimates is for most locations (except for Borgharen Dorp) lower than 1 m when looking at the average or median SAR-based water levels. In Lanaken, the average SAR water level deviates with just 1 cm from the LMW measurement, reaching within the 20 cm accuracy requirement. The remaining locations cannot achieve the 20 cm accuracy with respect to the LMW gauges. The standard deviation, or precision, on all location is less than 65 cm. The highest precision is computed at Bunde for the Sentinel-1 extent of 4 cm and at Lanaken for the Capella Space extent of 15 cm. The example at Lanaken clearly shows the differences between precision and accuracy. Despite the Capella Space data being twice as precise as the Sentinel-1 data, the accuracy of the Sentinel-1 estimates is higher in this scenario.

The location at Borgharen Dorp showed the highest differences for the Capella Space extent, and little to no detections by the Sentinel-1 outliers. The slope along the extent is approximately 13 degrees, which is in good agreement with the steep dikes that are surrounding the river. In addition, it can be seen that the dike is covered with grasses and some trees on both sides. The presence of trees and grasses can interfere with the signal, especially when dealing with X-band data.
The Sentinel-1 outlier detection only provided 1 overlapping point containing DEM-data along the edge of the extent. The height of that point is 41.28 m, which is over 2 m difference from the estimated height of the LMW location. The slope on this location is high, and the deviation of the DEM height within a 10 m radius from the edge point is 1.20 m. The high slope in the area can therefore lead to a variation of more than 1 m, within 10 m from the edge point. Considering the spatial accuracy of Sentinel-1, this suggests that 1 or 2 missing pixels in the outlier detection could result in an underestimation of the water level by more than 1 meter.

Comparison Sentinel-1 and Capella Space

In general, the validation of SAR to the LMW measurements has provided insights in the precision and accuracy of the SAR-based water levels. The precision of the SAR-based water levels for the entire AOI has been mentioned in the end of chapter 4: 0.141 m for Capella Space, 0.156 m for the Sentinel-1 T37 stack, and 0.210 m for the Combined Sentinel-1 extent. On a local level, the precision can vary more compared to the entire extent. As mentioned above, the standard deviation of the edge points is in all cases less than a meter. This indicates a high overall precision of all locations for both Sentinel-1 and Capella Space, that falls within the 20 cm range. The precision of Capella Space is only 0.014 m higher compared to Sentinel-1. The accuracy of the SAR-based water levels to the LMW measurements can be determined from the validation results. The 20 cm requirement set by Rijkswaterstaat cannot be consistently met when regarding the accuracy of the results but can be met when considering the precision of the results.

There are two factors that can contribute to precision and accuracy quality metrics: the slope and the density of vegetation within an area. A higher slope can have a large effect on an underestimated flood extent of the SAR data. As described above, for the location at Borgharen Dorp, the variation within 10 meters of an edge point can be 1.20 meters. For data sources with lower resolution pixels, one missed pixel may lead to an underestimation of more than a meter. The spatial resolution of Capella Space is 0.8 m, thus the slope is less of an issue in terms of pixels. However, the effects of under- or overestimation of the flood extent can still cause larger errors in the determination of the water level compared to flat areas.
Vegetation, such as grasses and trees, have an effect on the backscatter properties of the SAR signal. The misclassification resulting from interference of the signal translates into inaccuracies of the water levels. The influences of trees on X-band and C-band data makes the data less suitable to be used to measure floods that are beneath large canopies. The influences of trees on the SAR signal can be seen in Figure 5.7. The figure illustrates the dispersion of the signal that originates in trees and spreads across the river. Measuring floods with C-band SAR is more suitable for sparse vegetation compared to X-band, but L-band SAR is required for high vegetation with high penetration depth to the trunk of the trees.

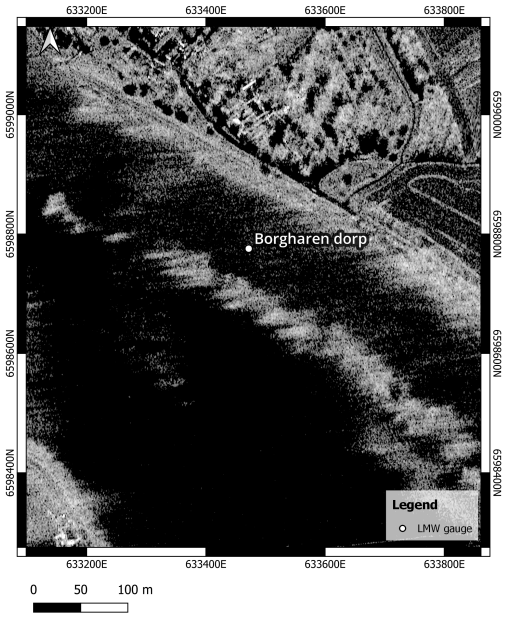


Figure 5.7: Capella Space SAR image (without speckle filtering) during the flood over the LMW location - Borgharen Dorp.

The key factors for comparing water level measurements from the LMW gauges to the SAR-based water levels, are summarized below:

- **Extent SAR & LMW gauges**: One of the main differences between the SAR-based water levels and the LMW gauge measurements is the location. The gauges measure the water level in the center of the river, and the SAR-based water levels are determined at the edges of the flood extent. The edges can be located close the gauges when considering small canals (Bunde and Borgharen Julianakanaal), but reaches further into the land for other areas (Lanaken).
- **Accuracy & Precision**: The accuracy of the SAR-based water levels is linked to the measurements from the LMW gauges. The gauges are treated in this thesis as the ”True” values. The accuracy of the SAR-based water levels (with respect to the LMW measurements) is higher for most locations than the 20 cm requirement. The precision, or standard deviation, of all the SAR edge points over the AOI is lower than 20 cm for both the Capella Space and Sentinel-1 data. The edge points around the LMW gauges range from 4 cm to 65 cm precision, as can be deduced from Table 4.9.
- **Overestimation and underestimation SAR**: Deviations from the SAR-based water levels to the LMW gauge measurements can have several reasons. Inaccurate computation of the flood extent translates into the water levels. Effects from vegetation have shown to result in underestimation of the flood extent, therefore underestimation of the water levels (Borgharen Dorp). Overestimation of SAR can be due to shadows or roads (mainly for the tresholding method) or puddle formations that do not represent the flooded area but are detected because of the presence of water. Lastly, steep slopes affect the water level estimation significantly. Both underestimation and overestimation of the flood extent can result into > 1 meter deviation in height.

5.2.2. Deltares Model

The SAR-based flood extents and Deltares model extent have been discussed in the previous section. In this section, the water level results of both sources will be discussed. The water level estimates have been visualized with the same color scale in Figure 4.23, Figure 4.30, Figure 4.32 and Figure 4.34. The Sentinel-1 T37 stack acquisition has been evaluated with two different moments of the model. The reason for this is that the model shows different results at 5:50 UTC compared to 07:00 UTC. Visually, the SAR results appear to align more closely with the results at 07:00 UTC. It is known that the timing of the model requires improvement. Based on the SAR-based flood extent, the timing of the model is off with approximately 1 hour and 10 minutes. Inflow and outflow of the water in the floodplain that can

be improved in further research. The combined flood inundation extent of the Sentinel-1 acquisitions is compared to the moment of the peak of the flood, which aligns with the acquisition of Capella Space on 15/07/2021 at 23:30 UTC.

The water levels of SAR and the model are composed differently. SAR-based water levels are determined on the edges of the flood extents, whereas the model interpolates the water heights over the entire extent. The comparison on the water levels will therefore only be performed on the edges of the flooded polygons derived from of the SAR data. The differences between the model results and the SAR data is plotted in Figure 4.24, Figure 4.31, Figure 4.33 and Figure 4.35 for Capella Space, Sentinel-T37 track with 5:50 UTC results, Sentinel-T37 track with 07:00 UTC results and Sentinel combined track with 23.30 UTC results, respectively.
Based on the distribution of the differences in water levels, the Capella Space flood extent shows most agreement of the model results within the accuracy range of -1 to 1m. Along the canals and the river edges on the North-West, some overestimation with respect to the model can be found. The differences in resolution between SAR and the model grids should be taken into account when interpreting these differences. When the edges of the grid cells or the SAR-flood extent are closely located to the canals, the slope can become steeper. Steep slopes can result in large variations of water levels on a small scale. This can affect the SAR-based water levels and deviate in a range of more than 1 meter.
Figure 5.8 displays the areas in more detail, underestimated areas are mostly located on places where the flood extent shows underestimation as well. Areas close to the river with an underestimation could be related to the presence of emergent vegetation, debris or high vegetation. Figure 5.8c is located South of Borgharen and is displayed and discussed in the previous section. Interference of the vegetation with the SAR signal (Figure 5.7) causes underestimation of the flood extent in the area which results in an underestimation of water levels.
Differences between SAR and the model results show an absolute accuracy within 1 meter over flat agricultural areas, locations with overestimation are mostly located along the canals. The locations along the canals indicate an underestimation of the water levels by the model.

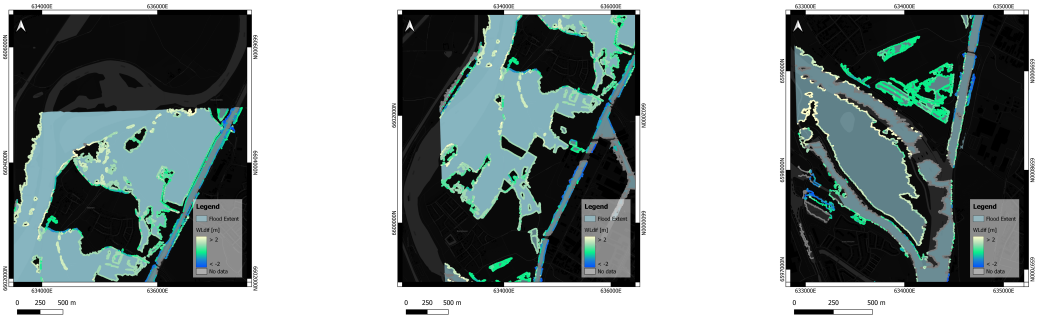
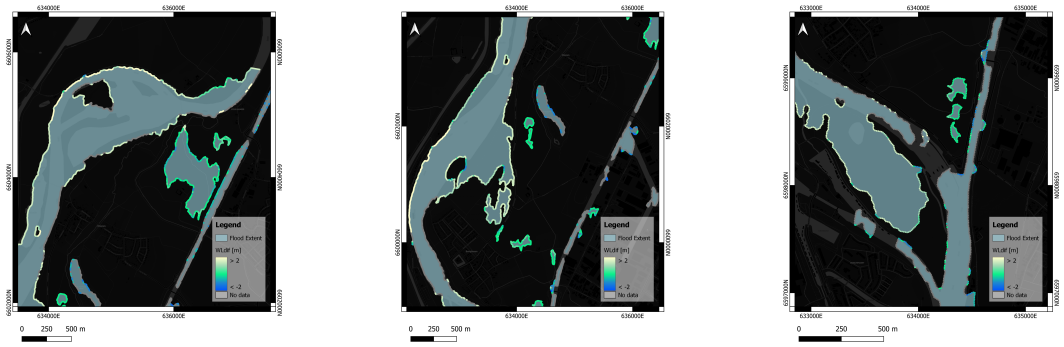


Figure 5.8: Zoom-in on the 3 largest areas of the AOI displaying the differences of between the model and Capella Space water levels at 23:30UTC: (a) Around villages of Bunde and Itteren (b) Between the villages of Itteren and Borgharen (b) South of Borgharen.

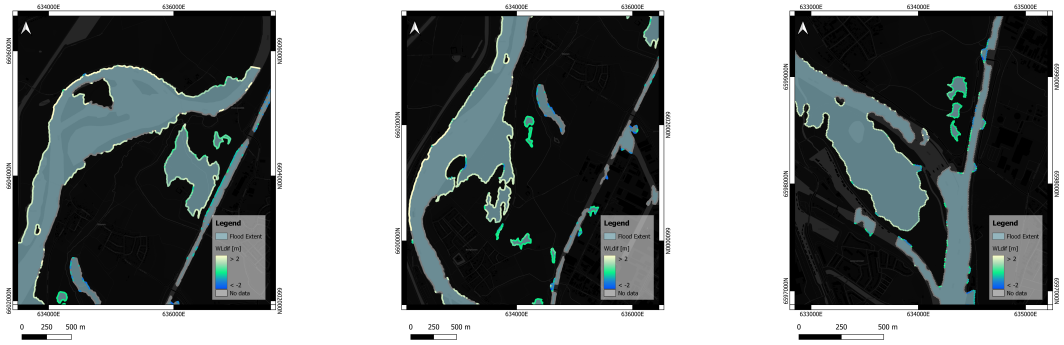
The differences between the model results and Sentinel-1 water levels, based on the T37 acquisition at 5:50 UTC, have been computed for two timestamps in the model. For both moments, the histogram of the values in Figure 4.31 and Figure 4.33 display that there is a peak of values that is underestimating the model results with approximately 2 meters. The differences in the areas ***b*** and ***c*** of Figure 5.9 and Figure 5.10 are minimal. For both results, underestimation of the water levels can be seen. The error matrices of the extents show approximately 64-68% True Positives, the remaining 32-36% of the False Negatives are partly related the underestimation as seen in these plots. Sentinel-1 shows more detail and variations along the edges of the flood extent, this might cause some areas to underestimate with respect to the larger and interpolated model grid cells.
The largest differences between the two model timestamps are shown in area ***a***. The model results at 05:50 UTC show an overestimation of the water levels. This is expected based on the visual interpretation that the Sentinel-1 flood extent is reaching further into the land compared to the model results.

The histograms of the differences between the two timestamps confirms this. The peak around -0.5 in the histogram of Figure 4.30, shifts to 0 in the histogram of Figure 4.32. This shift is related to the areas that were not yet flooded in the results at 5:50 UTC but were flooded in the results at 07:00 UTC. Generally, the results along the edges of the rivers and canals show an accuracy less than 1 m between SAR and model. Flat agricultural lands in the middle of the AOI also show similar water levels in both sources. Depending on the complexity of the location, SAR-based water levels of Sentinel-1 approximate the modeled water levels within a range of 1 m. For calibration purposes, utilization of data can be interpreted as input data, as it shows a higher agreement with the results at 07:00 UTC compared to those at 5:50 UTC.



(a) (b) (c)

Figure 5.9: Zoom-in on the 3 largest areas of the AOI displaying the differences of between the model and Sentinel-1 water levels at 5:50 UTC : (a) Around villages of Bunde and Itteren (b) Between the villages of Itteren and Borgharen (b) South of Borgharen.



(a) (b) (c)

Figure 5.10: Zoom-in on the 3 largest areas of the AOI displaying the differences of between the model at 07:00 UTC and Sentinel-1 water levels at 5:50 UTC : (a) Around villages of Bunde and Itteren (b) Between the villages of Itteren and Borgharen (b) South of Borgharen.

The differences between the model results and the combined results of the water levels, based on the Sentinel-1 T37 acquisition on 15^{th} of July and the T161 acquistion on 16^{th} of July, have been computed and shown in Figure 4.35. The combination of the two extents was done to generate the maximum flood inundation extent of the area. This extent relates to the peak of the flood, that is captured by Sentinel-1. The error matrices for both extents have been computed, the amount of True Positives is 46% for the combined Sentinel-1 flood inundation extent and 67% for Capella Space. This is an indication that the maximum extent of Sentinel-1 is underestimating the area that is computed from Capella Space. This is confirmed by the histogram of the difference between the model results and the Sentinel-1 extent. There is a peak of values that are underestimating the model results at approximately 3 m. There are some locations on the agricultural fields, edges of the rivers in the North and along the canals that show higher agreement with the model results within 1 m accuracy. The areas can be seen in Figure 5.11a All in all, the flood extent and water levels of the combined Sentinel-1 acquisitions shows some aligning

results when comparing to the Capella extent and Deltares model. However, a large part of the area is underestimating the extent and water levels on the corresponding locations. The agricultural areas located in the middle of the AOI, are a large part of the flooded area. These areas are not particularly prone to steep slopes or high vegetation and have been detected by the Capella Space acquisition. The reason that these areas are missing in the flood inundation extent of Sentinel-1 are related to the complexity of inundation detection. The image of the T161 track is taken two days after the peak of the flood. The maximum flood extent has retrieved and dried up before the second Sentinel-1 acquisition. Using the ATSA approach, the changes in soil moisture or presence of debris would have been detected when significant enough with respect to normal conditions.

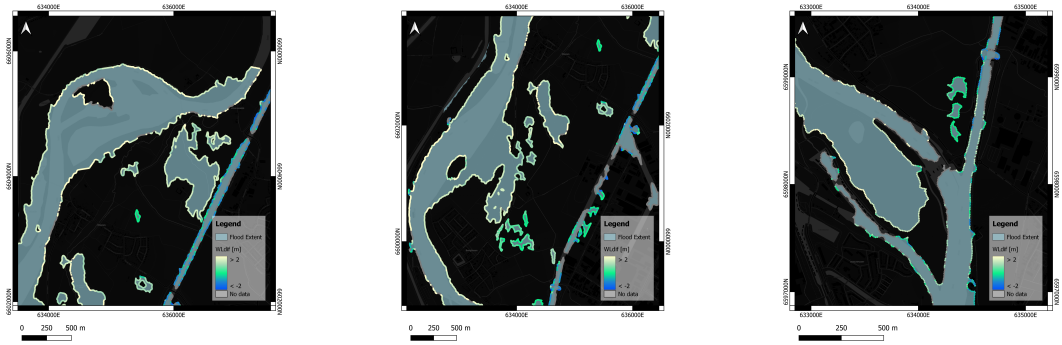


Figure 5.11: Zoom-in on the 3 largest areas of the AOI displaying the differences of between the model at 23:30 UTC and Sentinel-1 combined water levels of the two acquisitions at 15^{th} July and 17^{th} July: (a) Around villages of Bunde and Itteren (b) Between the villages of Itteren and Borgharen (b) South of Borgharen.

The key factors for comparing modeled results to SAR-based results, are summarized below:

- **Timing Model**: The results in Figure 4.13 indicate that the timing of the model is approximately 1 hour and 10 minutes off. The results from 07:00 UTC show a higher visual agreement with the SAR-based flood extent, compared to the results at 05:50 UTC. Using SAR imagery, the results of the model can be calibrated to the improve the timing of the inflow of the water. The results of the model can be aligned to the SAR-based flood extent or water levels at the acquisition time of the satellite.
- **Underestimation SAR**: As mentioned before, the presence of vegetation can affect the flood extent. Several areas throughout the AOI are showing underestimations of SAR due to disturbances of the signal. Small canals and steep slopes also have an effect on the flood extent, especially concerning the Sentinel-1 extent where one pixel difference can cause over 1 m change in water level. Underestimation of flood extent in these areas translate into inaccuracies of water level estimation.
- **Overestimation SAR**: There are several locations in the AOI that show distinct results between the model and SAR. Some results can be referred to underestimation of SAR due to presence of vegetation or overestimation from puddle formation detected by SAR. Other results are ambiguous, the first example is flooding of the dike that is not shown in the SAR images or aerial images, but is classified as flooded in the model. The second example are the fields in the South of the AOI, also shown in Figure 5.2, the fields are classified as flooded but the model does not classify the fields as flooded. The fields are either showing puddle formation or has been underestimated by the model. Both locations are shown in Figure 5.12, ground truth data can provide more insights into which data source is more accurate for these specific cases.
- **Extent SAR & Model**: The model covers the entire affected area along the Meuse River. Additional models exist to simulate flooding in the smaller streams as well, such as the Geul. Sentinel-1 SAR data can also provide simulations of the full area within a few days. However, Capella Space data covers a 5 km x 10 km area, which represents only a portion of either the river or the streams. This effect is displayed in the comparison between the data sources, the derived results from the Capella Space data are cut-off on the edges of the river.

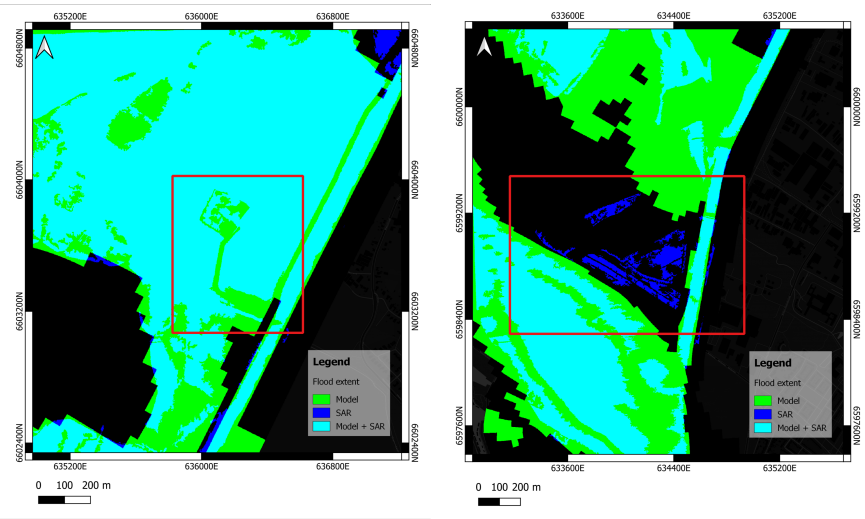
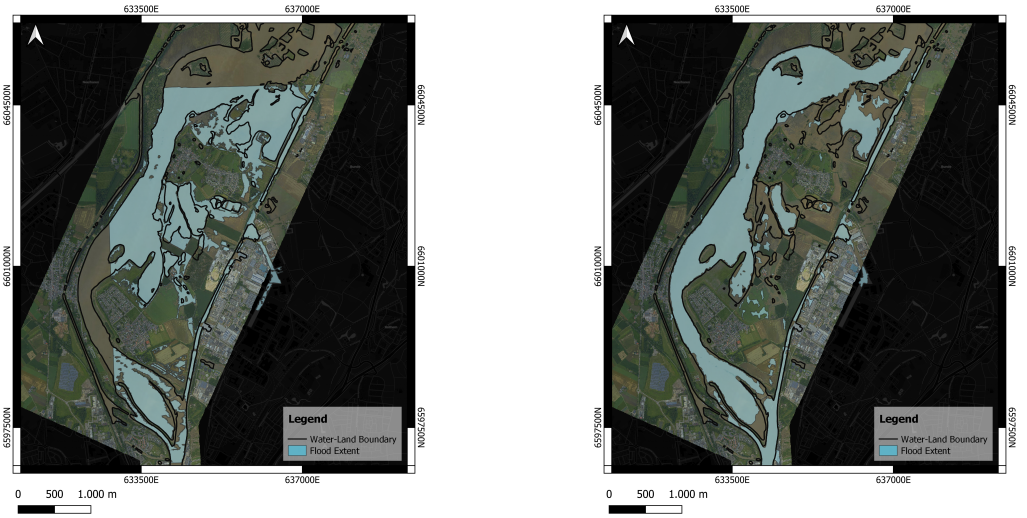


Figure 5.12: Locations where the modeled and SAR-based flood extent results display different classifications of flood water.

5.3. Aerial images and Water-Land Boundary lines

As mentioned in chapter 3, a visual comparison is performed as validation of the SAR-based flood extent. The aerial images are taken a few hours after the flood together with LiDAR measurements that have computed the Water-Land boundary lines. As both sources depict the maximum flood extent, the Capella Space flood extent and the combined Sentinel-1 extent of both acquisitions are used. Below an overview of the entire area is displayed to show the flight-lines of the planes with respect to the flood extents of the SAR satellites. Figure 5.14 zooms into a few locations of the AOI to depict the differences between the sources.
The aerial images confirm that the image acquisition of Capella Space was taken at the peak of the flood by correct classification for most of the flooded area. The aerial image overlaps most of the extent. The extent of the combined Sentinel-1 acquisitions and the Water-Land boundary lines both underestimate the maximum flood extent. The Sentinel-1 acquisitions have not been taken during the peak of the flood and therefore needs to rely on the detections made from inundation two days after the peak of the flood. The figures show that the extent reaches the boundaries close to the river and canal sides, but a few agricultural lands are not classified as flooded. The Water-Land boundary is derived from LiDAR measurements taken during the flight of the aerial images. It would be expected that these align to the extent provided in the aerial images. The derived extent from Capella Space shows that the lines are underestimating the maximum flood extent at all three locations.

One of the disadvantages of the Capella Space extent is that the image does not cover the entire AOI. The image edge lines are located across all three locations and therefore limit the estimated area to the image boundaries. Additionally, the interactions of dense vegetation with the SAR signal are clearly visible in the images. The location at image ***e***, shows that the distinctive fields containing dense vegetation are not classified, whereas aerial images show that floodwater is present. Similar phenomena can be seen at location ***h***, where the interference of the signal with the trees along the river results in less smooth the edges of the flood extent. This results in the underestimation of the extent that has also been discussed in the sections above.



(a) Capella Space flood extent.

(b) Sentinel-1 combined flood extent of the T37 and T161 acquisitions.

Figure 5.13: Overview of the SAR derived flood extents plotted on top of the aerial images the Water-Land boundary lines.

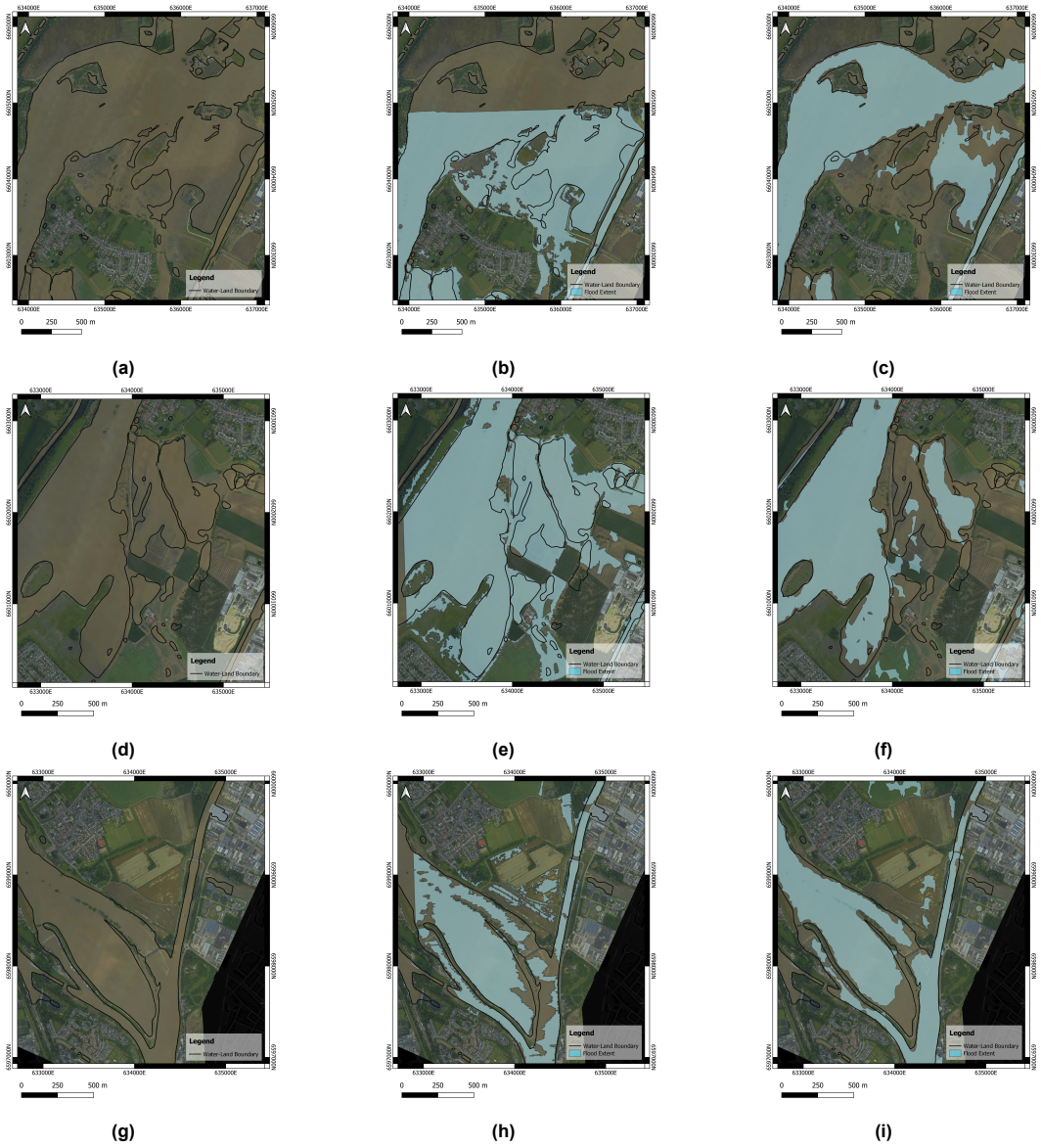


Figure 5.14: Zoom-in on 3 locations of the AOI displaying the differences between the Aerial images and Water-Land Boundary lines (a,d,g), to Capella Space flood extent (b,e,h), and the combined Sentinel-1 flood inundation extent (c,f,i).

5.4. Application of SAR during flood events

The aim of this thesis is to find the suitability of SAR-based flood mapping is for real-life applications, complementary to other resources. The capabilities of SAR are highly dependent on the application. Both SAR data sources have shown to detect approximately 64-68% of the flooded area (when compared to the model) and have a water level estimate with a precision of 0.141 m and 0.156 m, respectively for Capella Space and Sentinel-1 data. The flood extent derived from both Capella Space and Sentinel-1 contain False Positives and False Negatives. Some areas contain underestimation due to interferences of the radar signal with vegetation and other areas show overestimation containing large puddles detected as flooded water. On the other hand, the Capella Space image can map the flood extent over the demanded AOI within a few hours and Sentinel-1 flood extents can be generated over the entire river area within a few days.

The LMW measurements are known to have high accuracy but are only located on specific locations on the river. Along regional streams, some locations have no water level gauges present, or gauges can malfunction when the water reaches far out of the banks. The modeled data is known to have high accuracy when it comes to the maximum extent, but the details of the data can lack due to the use of

large grid cells and the implementation of winter vegetation instead of summer vegetation. The model data is also known to miscalculate the timing of the inflow and outflow of the water. An inaccurate timing of 1 hour and 10 minutes can be concluded from the SAR images. Satellite imagery can improve calibration of the flow of the water, alternatively to the gauges. SAR images cover the land, whereas the gauges are located along the river.
Another known advantage of using SAR for flood mapping is its independence of time of day and weather conditions. The time independency is of particular interest when the peak of the flood happens occurs during the night, as happened during the 2021 floods. Aerial images have to be taken by plane and, similar to optical satellites, are likely to be blocked by clouds that are accompanied with flood events. To generate the images, a flight plan has to be made and the weather conditions have to be right in order to takeoff. With the recent launch of Sentinel-1C, SAR images of Sentinel-1 are available four times per 6 days. The latest sub-meter resolution satellites from companies such as Capella Space, images are acquired on-demand and within a few hours.
Concluding, the use of SAR-based flood mapping should not be directly related to the same applications as the LMW gauges, nor the model data. The derived water levels are directly linked to the estimation of the flood extent which includes false positives and false negatives resulting in inaccuries of the water levels. Using SAR can improve calibration the timing of the models considering the flood extent over land. The SAR extent provided more information on water levels that are reaching outside the banks of the river compared to the gauge measurements. In case there is no other data available due to nighttime or cloud cover, SAR data can be a valuable tool to generate insights of the flood over the entire area.

6

Conclusion & Recommendations

6.1. Conclusion

In chapter 1, the topic and knowledge gaps were introduced following the research questions. This chapter summarizes the results and findings that were presented to answer the questions. First, the answers to the sub-research questions are addressed, which will then provide the foundation for answering the main research question.

6.1.1. Sub-questions

- *What are the most important components of the SAR signal that provide information about the flooded state of a pixel?*
  A SAR signal contains multiple components like amplitude, phase and polarization. Amplitude data is the most important and widely used component of the SAR signal when it comes to flood mapping. Water reacts with the signal as a specular reflection, meaning that the signal is scattered away from the satellite resulting in distinct low backscatter values. Utilization of phase information is of interest when dealing with changes of the coherent environment such as urban areas. Different polarizations are an important component when it comes to different land cover types. The amplitudes of different polarizations can enhance the delineation of the flood extent as both provide different information about the surface characteristics. Both coherence an polarimetry decrease the spatial resolution which makes it not suitable for this specific case study where high resolution is demanded for the estimation of water levels.

- *How can SAR imagery be combined with a Digital Elevation Model to accurately extract the water levels?*
  The results of the research have shown that an accurate estimation of the water levels is related to the flood extent. Deviation of a few meters in the (x,y) extent can lead to an inaccuracy in water level of a few meters. Flat areas have the advantage that the change of location in (x,y) will affect the estimation of the water level less rigorously. Combining the edges of the SAR-based flood extent with a DEM can provide insights in the longitudinal and transverse gradients of the flood extent. The accuracy of the water levels is dependent on the proper estimation of the flood extent and on the quality of the DEM.

- *How do the Water-Land Boundary estimation from on-demand sub-meter spatial resolution satellites compare to meter spatial resolution satellites with longer data archives?*
  The choice of data source is dependent on the purposes of the application. The sub-meter spatial resolution satellites provide more detail in smaller areas along the canals and dikes. Furthermore, the data can be provided on-demand which could be beneficial if the data is used to replace the acquisition of aerial images. The downsides of using this data is that there is usually one single image available, and coverage is dependent on the given AOI. Using a single image can lead

to falsely classified roads or tree shadows in the image as flooded. Another downside of using X-band data is the fact that it cannot penetrate through the canopy. If flooded area is located in a forest canopy, X-band is not a preferred data source. Using meter spatial resolution satellites with large archives has the advantage that it usually covers a large area, can estimate the trend over time, but the temporal resolution is lower leading to missed acquisitions during the peak of the flood. The results of the data are accurate enough for large flooded areas in open space but influences of high vegetated are also present in C-band SAR-systems. For more detailed information along dikes or canals, the data is too coarse in some locations. For calibration purposes of the flow and timing of the water, both data sources can be used.

- *How do the Water-Land Boundary estimations from SAR imagery compare to current resources?* LMW gauges can provide highly accurate water level measurement, however, the gauges are located on a few locations along the river. SAR measurements can capture the entire area at once and locate how far the water is reaching into the land. This information can be used to indicate the rising water in floodplains. Regarding model data, SAR imagery can provide more insights for the timing and flow of the water. Model data can estimate the flood extent with high accuracy, but the timing of the models contain offsets of the real flood developments. In addition, the results of the model data contain errors due to incorporation of winter vegetation instead of summer vegetation. For the purpose of extent mapping, aerial images can also be used. Aerial images can show with high resolution the extent of the water over the area. During a flood event, these images are taken by plane. In case that weather conditions do not allow for plane takeoff, or when the peak of the flood occurs at night, SAR systems can provide data. Comparing the SAR-based results to the current Water-Land boundary lines of the maximum extent, there was a higher agreement to the aerial images than to the lines. Indicating that SAR imagery of Capella Space has a higher accuracy compared to the LiDAR derived Water-Land boundary lines.

6.1.2. Final Conclusions

The main research question of this thesis is:

What is the value of SAR-based flood mapping as a complement to flood models, in-situ measurements and other remote sensing data for flood events in the Netherlands?

The results of the research have shown that SAR-based flood mapping contains complementary information to current resources. The required water level accuracy set by Rijkswaterstaat is 20 cm. When comparing the results of the SAR-based water levels to the LMW gauge measurements, this accuracy cannot be consistently achieved. However, the precision of the results falls within the 20 cm range for several locations and the overall precision. SAR data can cover a broader area compared to specifically located LMW gauges. Capella Space provides a coverage of about 5 km x 10 km, while Sentinel-1 covers the entire country within a few days. The model data also spans the full river area, but the results indicate that the timing of the water inflow does not align with the SAR-based results. The modeled results appear to be off by approximately 1 hour and 10 minutes.
. Using a conventional thresholding method for the sub-meter resolution satellite data of Capella Space provided a flood extent that contained some false positives in areas such as airport runways or shadows of trees. Additionally, the X-band wavelength of the satellite caused locations of dense vegetation to be prone to volume scattering. This led to underestimation of the flood extent at some locations, leading to inaccuracies in the estimates of the water levels. Most of the areas showed a precision of <65 cm and <1 m accuracy compared to the LMW gauges. For applications such as real-time flood extent mapping or mapping along small canals with dikes, the sub-meter resolution is necessary in order to approach an accurate water level estimate based on the Sentinel-1 data providing limited data on these locations.
The results of Sentinel-1 used an Amplitude Time Series Analysis (ATSA) to detect outlier pixels related to floods. The outliers formed the maximum flood extent of two dates. Comparing the extent and estimated water levels to the model data and aerial images showed a large underestimation in both extent and height estimates. From these results it can be concluded that in order to capture the maximum flood extent with Sentinel-1, the satellite acquisition has to be closer to the peak of the flood.
Instead of looking at the combined extent, the extent and water levels derived from the T37 stack (acquisition at 15/07/2021 5:50UTC), provided insights about the timing of the model data being ap-

proximately 1 hour and 10 minutes off. Comparing the SAR-based water levels to the LMW gauges showed a precision of <60 cm and <1 m accuracy compared to the LMW gauges, which is similar to the results of water levels derived by Capella Space. The advantage of the approach is that the results can be generated over the entire river area and for each acquisition date in the time series. The main applications for this data therefore lie in the potentials of large scale flood mapping and getting an indication of the flood extent at the specific acquisition times. The extents are agreeing on various locations with the model and LMW data but high vegetation and slopes lead to wrongly classified flooded areas. One to two misclassified pixels can lead to >1 m difference in water level.
Concluding, the complementary information of SAR are the large scale implementation (Sentinel-1), flexible AOI acquisitions (Capella Space), possibilities of acquisitions during the night or during cloudy weather on-demand (Capella Space) or consistently over time (Sentinel-1), and coverage over land instead of river measurements.

The results of this thesis have been purely based on SAR-data as the goal was to investigate this data type, while limiting the amount of assumptions. Below a few recommendations are mentioned that could improve the results for future research.

6.2. Recommendations

6.2.1. Flood extent

The thresholding method used for Capella space is a robust, simple, yet effective approach. A few optimizations are recommended to further explore this method and improve the amount of correct classifications. Further research is needed to test the improvements of the classifications using a baseline image. Using the baseline image, the difference between the flooded image and the baseline can be computed. The usage of a baseline image can reduce the amount of false positives on locations. In addition, more advanced segmentation methods can be evaluated such as Gaussian Mixture Models that have shown promising results [46].

Another way to improve the results is the implementation of region growing. Region growing is the process of using the underlying DEM to classify pixels that are located close to flooded pixels but are classified as non-flooded. The change in classification will be implemented if the flooded pixels have a lower DEM values compared to the flooded neighboring pixels. Using this approach, the flood extent can be enlarged using the likeliness of pixels being flooded. This could similarly be implemented for the flood extent derived from Capella Space and Sentinel-1.

Vegetation shown to influence the estimation of the flood extent and thereby the water levels. This is due to the interference of vegetation on the signal. The launch of S-band and L-band satellites could improve the results along densely vegetated areas as they have a larger penetration depth. Research into the applications of the NISAR satellite that will have a spatial resolution of approximately 3-48 meters (depending on the mode) is therefore recommended [47].

6.2.2. ATSA

The Sentinel-1 Amplitude Time Series Analysis can be improved in several ways. Using a similar approach but with different polarization or coherence data, more outliers can be detected in the area. Polarizations can help when dealing with different types of landcover. The surface characteristics may vary and interact differently with VV or VH polarization. Likewise, in the more complex urban environments different polarizations may generate distinct backscatter patterns [27]. This may also be visible in the coherence data over an area. The coherence of the signal is a measure of interferometric correlation and it expected to decrease for flooded areas between two acquisitions. To complement the amplitude data more studies are using coherence data for flood mapping in urban areas. The downside of coherence data is the drop in spatial resolution. If the loss of spatial resolution can be permitted within a chosen AOI, the implementation of coherence data into the ATSA approach is recommended.

Despeckling of the SAR data has been briefly mentioned in the background but has been used in conventional ways in the thesis. For further research the implementation of more advanced speckle filters such as the MERLIN filter could improve the ATSA results. By using a more advanced speckle filtering, the time series can become smoother. Time series containing less noise are expected to have a more defined dispersion of the data, which could lead to an improved outlier detection.

6.2.3. Model input
The integration SAR-based flood extents into the models can be further evaluated. The focus has been on the edges of the flood extent, which presents challenges in certain areas. It is estimated and known that the models encounter difficulties with the timing and phase of the flood. Consequently, it can be expected that combining the strengths of observations and models can generate an improved result. Proper incorporation of the SAR data into the model is beyond the scope of this thesis, but is recommended to further developments.

6.2.4. LMW Validation
The approach for both data sources is applied to the large Meuse river in the selected area. Sentinel-1 covers the entire river and could be validated to other LMW locations or gauges from the Waterboards along the streams. Nonetheless, villages and cities along the streams, such as the Geul, have been excluded from this research. During the flood event, the city of Valkenburg was severely affected by the water coming from the Geul stream. Future research focusing on the AOI including locations and/or cities along the streams can investigate the application of the approach to urban areas and smaller water bodies. The complexity of the urban area can affect the results or applicability of the approach.

6.2.5. Scaling
Lastly, for all recommendations regarding this subject. It is important to keep in mind the effectivity of the approach. Flooding is a phenomena that occurs along the rivers and streams, throughout the entire country. The implementation should be scalable and reproducible to other areas or other flood events.

References

[1] P.A. Arias et al. “Technical Summary”. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2021, 33−144. DOI: 10.1017/9781009157896.002.

[2] S.I. Seneviratne et al. “Weather and Climate Extreme Events in a Changing Climate Supplementary Material”. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. 2021. URL: Available%20from%20https://www.ipcc.ch/.

[3] Zurich. *Three common types of floods explained*. https://www.zurich.com/knowledge/topics/flood-and-water-damage/three-common-types-of-flood. [Accessed 22-12-2024].

[4] Bernice Rosenzweig et al. *NPCC4: Climate change and New York City’s flood risk*. Tech. rep. Wiley Online Library, 2024.

[5] Planbureau voor de Leefomgeving. *Kleine kansen, grote gevolgen*. Accessed 17-08-2023. URL: https://themasites.pbl.nl/o/risico-overstromingen/#:~:text=Ongeveer%2059%25%20van%20het%20Nederlandse,rampen%20die%20Nederland%20kunnen%20treffen..

[6] Abdella Salem and Leila Hashemi-Beni. “Inundated Vegetation Mapping Using SAR Data: A Comparison of Polarization Configurations of UAVSAR L-Band and Sentinel C-Band”. In: *Remote Sensing* 14.24 (2022), p. 6374.

[7] Kennisportaal Klimaatadaptatie. *How is precipitation changing?* https://klimaatadaptatienederland.nl/en/knowledge-dossiers/themes/waterlogging/how-precipitation-changing/. [Accessed 14-12-2024].

[8] Task Force Fact Finding Hoogwater et al. “Hoogwater 2021 Feiten en Duiding”. In: *ENW report* (2021).

[9] Het Waterschapshuis. *Geodata Hoogwater limburg*. https://storymaps.arcgis.com/stories/7488a4903f47499a9f765a23619eb2f4. [Accessed 27-10-2024]. 2022.

[10] Copernicus. *Copernicus Emergency Management Service - Mapping*. Accessed 14-12-2024. URL: https://emergency.copernicus.eu/mapping/#zoom=2&lat=13.56036&lon=33.82273&layers=0BT00.

[11] GISGeography. *The Atmospheric Window in Remote Sensing - GIS Geography — gisgeography.com*. https://gisgeography.com/atmospheric-window/. [Accessed 21-04-2024].

[12] Capella Space. *Imagery*. https://www.capellaspace.com/products/imagery/. [Accessed 22-08-2024].

[13] SentiWiki. *S1 Mission*. https://sentiwiki.copernicus.eu/web/s1-mission. [Accessed 07-02-2025].

[14] Earthdata. *What is Synthetic Aperature Radar?* https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar. [Accessed 22-08-2024].

[15] Maria F Lopez Ornelas. “The Mexican Water Forest: benefits of using remote sensing techniques to assess changes in land use and land cover”. In: (2016).

[16] Carlos López Martinez. *Multidimensional speckle noise. Modelling and filtering related to sar data.* Universitat Politècnica de Catalunya, 2003.

[17] Frederik Johannes Van Leijen. “Persistent scatterer interferometry based on geodetic estimation theory”. In: (2014).

[18] EO College. *Echoes in Space*. https://eo-college.org/courses/echoes-in-space/. [Accessed 22-08-2024].

[19] Jet Propulsion Laboratory. *Polarimetry*. https://nisar.jpl.nasa.gov/mission/get-to-know-sar/polarimetry/. [Accessed 22-08-2024].

[20] Weiying Zhao et al. “RABASAR: A fast ratio based multi-temporal SAR despeckling”. In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE. 2018, pp. 4197–4200.

[21] Aiyeola Sikiru Yommy, Rongke Liu, and Shuang Wu. “SAR image despeckling using refined Lee filter”. In: *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*. Vol. 2. IEEE. 2015, pp. 260–265.

[22] Rui Zhang et al. “A novel feature-level fusion framework using optical and SAR remote sensing images for land use/land cover (LULC) classification in cloudy mountainous area”. In: *Applied Sciences* 10.8 (2020), p. 2928.

[23] Francesco Lattari et al. “Deep learning for SAR image despeckling”. In: *Remote Sensing* 11.13 (2019), p. 1532.

[24] Emanuele Dalsasso, Loïc Denis, and Florence Tupin. “As if by magic: self-supervised training of deep despeckling networks with MERLIN”. In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021), pp. 1–13.

[25] Fatemeh Foroughnia et al. “Evaluation of SAR and optical data for flood delineation using supervised and unsupervised classification”. In: *Remote Sensing* 14.15 (2022), p. 3718.

[26] Michael Andrew Manalili. “Flood Delineation of Synthetic Aperture Radar data using python”. In: *University of The Philippines* (2018).

[27] Ramona Pelich et al. “Mapping floods in urban areas from dual-polarization InSAR coherence data”. In: *IEEE Geoscience and Remote Sensing Letters* 19 (2021), pp. 1–5.

[28] David C Mason et al. “Detection of flooded urban areas in high resolution Synthetic Aperture Radar images using double scattering”. In: *International Journal of Applied Earth Observation and Geoinformation* 28 (2014), pp. 150–159.

[29] Donato Amitrano et al. “Flood detection with SAR: A review of techniques and datasets”. In: *Remote Sensing* 16.4 (2024), p. 656.

[30] Marco Chini et al. “SAR coherence and polarimetric information for improving flood mapping”. In: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE. 2016, pp. 7577–7580.

[31] Rosa Colacicco et al. “High-Resolution Flood Monitoring Based on Advanced Statistical Modeling of Sentinel-1 Multi-Temporal Stacks”. In: *Remote Sensing* 16.2 (2024), p. 294.

[32] eoPortal. *Copernicus: Sentinel-1*. https://www.eoportal.org/satellite-missions/copernicus-sentinel-1#copernicus-sentinel-1--the-sar-imaging-constellation-for-land-and-ocean-services/. [Accessed 22-08-2024].

[33] Celestrak. *SATCAT*. https://celestrak.org/satcat/table-satcat.php?NAME=capella&PAYLOAD=1&MAX=500. [Accessed 22-12-2024].

[34] European Space Agency. *Mission ends for Copernicus Sentinel-1B satellite*. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1/Mission_ends_for_Copernicus_Sentinel-1B_satellite. [Accessed 22-08-2024].

[35] Actueel Hoogtebestand Nederland. *AHN 4*. https://www.ahn.nl/ahn-4. [Accessed 12-02-2025].

[36] Earth Data. *Digital Elevation/Terrain Model (DEM)*. https://www.earthdata.nasa.gov/topics/land-surface/digital-elevation-terrain-model-dem. [Accessed 09-02-2025].

[37] Agentschap Digitaal Vlaanderen. *Digitaal Hoogtemodel Vlaanderen II*. https://www.vlaanderen.be/digitaal-vlaanderen/onze-diensten-en-platformen/earth-observation-data-science-eodas/het-digitaal-hoogtemodel/digitaal-hoogtemodel-vlaanderen-ii. [Accessed 10-02-2025].

[38] WL Rapporten. “Over referentievlakken en peilschaalopmetingen in het Zeescheldebekken”. In: (2015).

[39] Rijkswaterstaat. *Meten bij Rijkswaterstaat*. 2024. URL: https://www.rijkswaterstaat.nl/water/waterbeheer/metingen/meten-bij-rijkswaterstaat.

[40] Panchagnula Manjusree et al. “Optimization of threshold ranges for rapid flood inundation mapping by evaluating backscatter profiles of high incidence angle SAR images”. In: *International Journal of Disaster Risk Science* 3 (2012), pp. 113–122.

[41] Nobuyuki Otsu et al. “A threshold selection method from gray-level histograms”. In: *Automatica* 11.285-296 (1975), pp. 23–27.

[42] Khuong H Tran, Massimo Menenti, and Li Jia. “Surface water mapping and flood monitoring in the Mekong Delta using sentinel-1 SAR time series and Otsu threshold”. In: *Remote Sensing* 14.22 (2022), p. 5721.

[43] Mark Edwin Tupas et al. “An intercomparison of Sentinel-1 based change detection algorithms for flood mapping”. In: *Remote Sensing* 15.5 (2023), p. 1200.

[44] Francisco Carreno Conde and Maria De Mata Munoz. “Flood monitoring based on the study of Sentinel-1 SAR images: The Ebro River case study”. In: *Water* 11.12 (2019), p. 2454.

[45] Encord. *Confusion Matrix*. https://encord.com/glossary/confusion-matrix/. [Accessed 27-10-2024].

[46] Nestor Yague-Martinez et al. “Towards frequent flood mapping with the Capella SAR system. The 2021 Eastern Australia floods case”. In: *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE. 2021, pp. 6174–6177.

[47] eoPortal. *NISAR (NASA-ISRO Synthetic Aperture Radar)*. https://www.eoportal.org/satellite-missions/nisar#mission-capabilities. [Accessed 07-12-2024].