UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

# Integrated system for detecting algal blooms in the Arabian/Persian Gulf: A multi-sensor approach using the Sentinel-2 and Sentinel-3 satellites data


| | |
|---|---|
| Author | Thonnard Julien |
| Supervisor | Hanert Emmanuel (ELIE) |
| Readers | Delhez Baptiste (ELIE) |
| | Wang Qiang (ELIE) |
| Academin year | 2023 - 2024 |


Thesis presented in partial fulfilment of the requirement for the degree of Master in Bioengineering : Environmental Bioengineering

# Acknowledgements

I would like to start by expressing my gratitude to Emmanuel Hanert for his guidance, kindness, time, and interest in my work. His continuous availability and support, as well as his role as the main point of contact for collaborations with other researchers and institutes, have been invaluable.

I would also like to thank Ponnumoy Vethamony of Qatar University for the references he sent us, and Yousef Alosairi of KISR for exchanging data and results. Many thanks to them for their reactivity and their time.

I would also like to thank the SLIM team for their advice and guidance throughout this year. A special thanks to Alexis, who followed me closely as a PhD student but also as a friend, for his advice and ideas, as well as his time and patience.

I would also like to thank all the people at the ELIE department who have helped me get through this year in the best possible conditions.

Then, a big thank you to the student team for sharing this year together and supporting each other throughout.

I would like to thank all my friends for their support and interest during this period. A special mention to Basile and Mattias for being there for each other in the final stages, and heartfelt thanks to Lola for her encouragement and assistance with my thesis.

Finally, a huge thank you to my family for supporting me throughout this year and my studies. You have enabled me to reach my goals and become the person I am today.

# CONTENTS


**1 Introduction** **1**

1.1 Algal blooms . . . 1

1.1.1 Characteristics . . . 1

1.1.2 Causes . . . 3

1.1.3 Consequences . . . 4

1.1.4 Development . . . 5

1.1.5 Dynamic . . . 5

1.1.6 Historical events . . . 7

1.2 Detection of algal blooms . . . 8

1.2.1 Remote sensing . . . 9

1.2.2 Chlorophyll . . . 10

1.2.3 ERGB images . . . 10

1.2.4 Detection index . . . 10

**2 Objectives** **11**

**3 Materials and methods** **12**

3.1 Zone of interest . . . 12

3.1.1 Climate . . . 13

3.1.2 Hydrographic situation . . . 13

3.1.3 Current geopolitical situation . . . 15

3.2 Algal blooms detection . . . 16

3.2.1 Satellites imagery . . . 16

3.2.2 Model calibration images . . . 18

3.2.3 Sentinel-3 initial quick check for blooms . . . 19

3.2.4 Sentinel-2 process . . . 24

3.2.5 Algal bloom visualisation . . . 28

3.2.6 Method summary . . . 29

3.3 Validation . . . 30

3.3.1 Sentinel-3 . . . 30

3.3.2 Sentinel-2 . . . 30

**4 Results** **33**
4.1 Model calibration . . . 33
4.1.1 Sentinel-3 . . . 33
4.1.2 Sentinel-2 . . . 35
4.2 Model limitation . . . 40
4.2.1 Sentinel-3 . . . 41
4.2.2 Sentinel-2 . . . 42
4.3 Trend over 2023 . . . 43
4.3.1 Sentinel-3 . . . 43
4.3.2 Sentinel-2 . . . 46
4.4 Link with oil spills . . . 48

**5 Discussion** **49**

**6 Conclusion** **53**

**Bibliography** **i**

**A Table of images without algal blooms used for validation** **vi**

# LIST OF FIGURES


1.1 Examples of different algal blooms (on the left in Baltic Sea and on the right in Barents Sea ("Image of the day | Copernicus", n.d.)). . . . . 2

1.2 Representation of an algal bloom front (Anderson, 2001). . . . . 6

1.3 Number of recorded HABs for 12 distinct geographical regions. (ANZ = Australia New Zeland. BENG = Benguela Current. CCA = Central America Caribbean. ECA = East Coast America. EUR = Europe. IND = Indian Ocean. MED = Mediterranean. NAS = North Asia. PAC = Pacific. SAM = South America. SEA = South East Asia. WCA = West Coast America) (Hallegraeff, 2021). . . . 7

1.4 Range of spectral bands for certain optical satellites (Radočaj *et al.*, 2020) . . . 9

3.1 Geography of the Gulf with the representation of different marine surface currents in the Gulf, with blue arrows for low-salinity flows, red arrows for high-salinity flows and purple arrows for transition zones (from Michael Reynolds, 1993; Vaughan *et al.*, 2019). . . . . 12

3.2 Representation of summer and winter marine surface currents in the Indian Ocean (Shankar *et al.*, 2002). . . . . 14

3.3 Enhanced RGB image based on Sentinel-2 bands 08, 04 and 03 showing algal blooms in the Gulf (left) and in the Pyramid Lake (right). . . . . 18

3.4 Example of an image before (left) and after (right) the application of the masks shown in Table 3.2. . . . . 20

3.5 Comparison between the different way to compute OTCI (blue represents zero chlorophyll, green represents low chlorophyll (water), red represents high chlorophyll, and black represents no chlorophyll). . . . . 22

3.6 Representation of the masked area in hatched red and the area under consideration in red. . . . . 23

3.7 Comparison of pixels highlighted (yellow) at each stage of analysis for the image of 04 may 2019 in the Gulf of Oman, with the red arrow indicating the presence of a ship. . . . . 27

3.8 Representation of the buffer with 10 iterations (left) and of the result of the intersection between the buffer and the ABDI layer (right) for the image of 04 may 2019 in the Gulf of Oman. . . . . 28

3.9 General detection scheme with brown boxes corresponding to Sentinel-3 processes and blue ones corresponding to Sentinel-2 processes. . . . . 29
3.10 Representation of the general appearance of a confusion matrix. . . . . 30

4.1 Trend in accuracy for each parameter set (ABDI - NIR - ABDI_bis) as a function of buffer value. Solid, dashed and dotted lines respectively represent cases with ABDI_bis = 0, 1 and 2. . . . . 36
4.2 Comparison of the evolution of precision (orange), recall (purple) and F1-score (red) as a function of buffer value for each set of parameters (ABDI - NIR - ABDI_bis). . . . . 37
4.3 ROC curves as a function of different buffer sizes for each parameter set (ABDI - NIR - ABDI_bis). . . . . 39
4.4 Comparison of surface chlorophyll concentration (green) and turbidity (red) for buoy KISR01 (29.436000° 47.903553°) in central Kuwait Bay. . . . . 40
4.5 Comparison of the number of pixels assigned to the algal bloom category according to two zones (the whole Bay in hatched and the Bay without the northern zone in red). . . . . 41
4.6 Comparison of the chlorophyll concentration at the buoy (purple) and the OTCI value for the pixel containing the buoy (orange). . . . . 42
4.7 On the left, a comparison of the predicted algal bloom zone (purple) and the actual algal bloom zone (hatched). On the right, comparison between the algal bloom zone (red) and the calculated ABDI (the greener the more likely it is to be an algal bloom). . . . . 42
4.8 Number of pixels whose value is greater than or equal to 2.5 for OTCI month by month for Sentinel-3A (in orange) and Sentinel-3B (in purple). The threshold value of pixels that must be exceeded to predict an algal bloom is shown with a dashed line. . . . . 43
4.9 Number of days with algal blooms for each month of 2023 for the broad area including the Gulf, the Strait of Hormuz and the western Gulf of Oman (left panel) and only for the Gulf (right panel) . . . . . 44
4.10 Map showing the areas most affected by algal blooms in 2023 and the location of desalination plants in black dots. . . . . 45
4.11 Delineation of the February 17, 2023 algal bloom (red) using the Sentinel-2 image method. The last close-up view corresponds to the extent of the image shown in Figure 4.12. . . . . 46
4.12 Green-spectrum enhanced RGB image of the algal bloom of February 17, 2023. This view corresponds to the last close-up view shown in Figure 4.11. . . . . 47
4.13 Comparison between the oil spill of March 28, 2021 (black) and its extension to April 1, 2021 (red) and the algal bloom observed on April 1, 2021 (green). . . . 48

# LIST OF TABLES

1.1 Summary of the main causes of algal blooms (Zohdi and Abbaspour, 2019). . . . 3

3.1 Summary of satellite images used for the validation process with the presence of algal blooms. . . . 18
3.2 Summary of mask layers used . . . 20
3.3 Description of the different classes in the SCL map ("Preprocessing Sentinel-2 images — eo-toolbox 0.1 documentation", n.d.) . . . 25
3.4 Comparison of the number and proportion of pixels highlighted at each stage of analysis for the image of 04 may 2019 in the Gulf of Oman. . . . 27

4.1 Validation results for pixels with OTCI value $\geq 1.8$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image. . . . 34
4.2 Validation results for pixels with OTCI value $\geq 2.5$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image. . . . 34
4.3 Validation results for pixels with OTCI value $\geq 4$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image. . . . 35
4.4 Summary of value ranges taken into account for model calibration of Sentinel-2 images . . . 35
4.5 Highest F1-score for each hyperparameter set (ABDI - NIR - ABDI_bis) and its associated buffer value. . . . 38

A.1 Summary of satellite images used for the validation process without algal blooms. vi

# List of abbreviations and acronyms

**ABDI** Algae Blooms Detection Index

**ERGB** Enhanced RGB

**ESA** European Space Agency

**FAI** Floatting Algae Index

**FN** False Negative

**FP** False Positive

**FPR** False Positive Rate

**FR** Full Resolution

**HAB** Harmful Algal Bloom

**NDVI** Normalized Difference Vegetation Index

**NIR** Near-InfraRed

**OLCI** Ocean and Land Colour Instrument

**OTCI** OLCI Terrestrial Chlorophyll Index

**RGB** Red - Green - Blue

**ROC** Receiver Operating Characteristic

**ROPME** Regional Organisation for the Protection of the Marine Environment

**SCL** Scene CLassification

**SNAP** SeNtinel Application Platform

**TN** True Negative

**TP** True Positive

**TPR** True Positive Rate

**WGS** World Geodetic System

CHAPTER

# ONE

# INTRODUCTION

## 1.1 Algal blooms

### 1.1.1 Characteristics

Each aquatic environment abounds with plant and animal life, including over 5,000 species of microscopic algae, also known as phytoplankton (Zohdi and Abbaspour, 2019), spanning from single-celled to multicellular organisms. When these algae undergo rapid proliferation, it gives rise to a phenomenon referred to as a bloom (Anderson *et al.*, 2017). Approximately 300 of these bloom species can result in dense and uncontrollable growth, posing a challenge for aquatic ecosystems (Zohdi and Abbaspour, 2019). These events can be observed in any marine environment, in a river, lake or sea. Moreover, their location is often correlated with temporal and size variables. On one hand, for example, algal blooms along coastlines or covering small areas will often be temporary, lasting from a few days to a few weeks. On the other hand, when these events are large enough to be visible on satellite images, more than several kilometers, they last longer, ranging from several weeks to a whole year (Anderson *et al.*, 2017), such as the one that occurred between 2008 and 2009 in the Gulf of Oman and the Arabian/Persian Gulf (Richlen *et al.*, 2010; Zhao *et al.*, 2015; Yu *et al.*, n.d.). The latter are more often observed in seas and oceans, and even along coastlines (Anderson *et al.*, 2017).

An additional crucial feature of algal blooms is their color (Figure 1.1). The colour is determined by that of the pigments or other cellular components in the algae that proliferate (Schleyer and Vardi, 2020). Indeed, each algal species responsible for blooms is associated with a specific color, which can range from red and green to other hues like orange or brown. However, if the concentration of algae does not reach a threshold, the characteristic color may not be visible. Nevertheless, even colorless algal blooms are considered as algal blooms as color is not the only parameter to distinguish blooms from water (Anderson *et al.*, 2017).

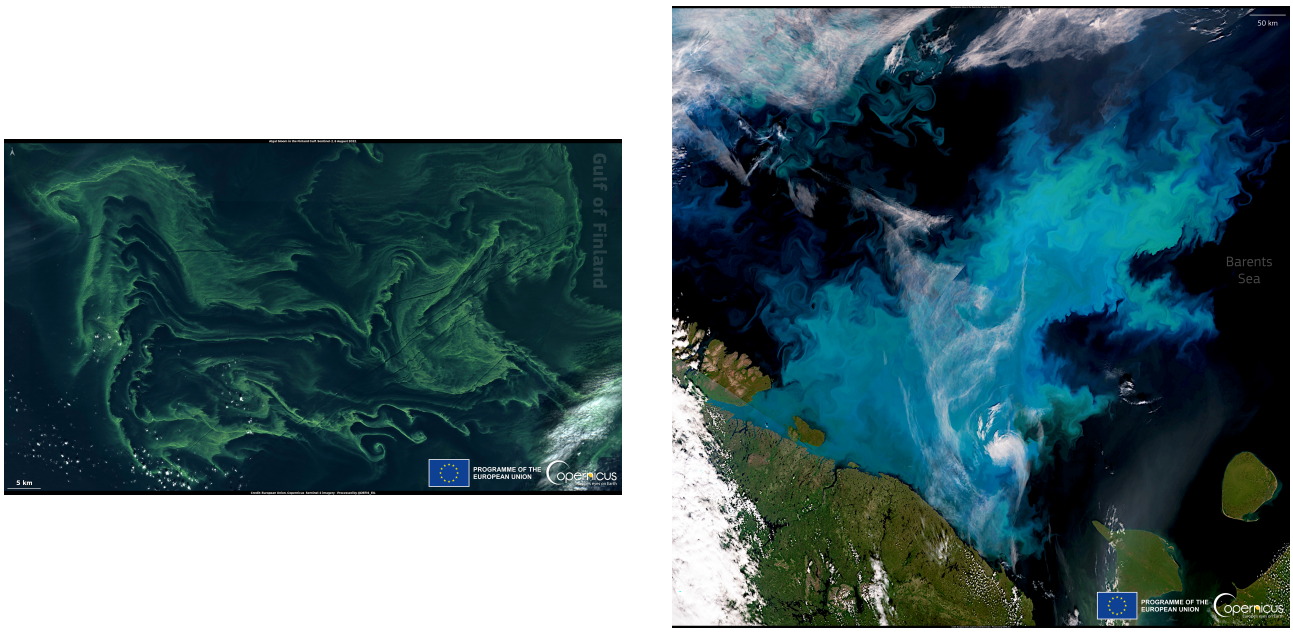



Figure 1.1: Examples of different algal blooms (on the left in Baltic Sea and on the right in Barents Sea ("Image of the day | Copernicus", n.d.)).

Algal blooms are divided into two categories: non-harmful and harmful. The latter are refered to as Harmful Algal Blooms or HABs. They are triggered by unicellular microscopic algae called phytoplankton (Anderson *et al.*, 2017). They encompass approximately 75 species of phytoplankton, such as dinoflagellates, diatoms, *Haptophyceae*, *Cyanophyceae*, and some cycloflagellates (Zohdi and Abbaspour, 2019). Nevertheless, HABs are not solely induced by microscopic algae; macroalgae can also serve as the origin. In fact, those are on the rise and persist longer than those triggered by phytoplankton (Anderson *et al.*, 2017). This is due to changes in the availability of phosphorus and nitrogen, as well as changes in the hydrographic situation caused by human activities. The first areas to be affected by these changes are the shallow coastal zones impacted by tides. It is in these areas that macroalgae are found, and it is these algae that are the first to suffer from these disturbances (Ansell *et al.*, 1998). Moreover, there are two categories of HABs: those that are toxic and those that are non-toxic (Anderson *et al.*, 2017). A toxic HAB is identified by the generation of potent toxins (Zohdi and Abbaspour, 2019).

However, it is important to point out that, generally speaking, algae and, by extension, algal blooms have biogeochemical characteristics that are necessary for several natural cycles. Indeed, algal blooms influence the cycles of elements such as carbon, sulphur and nitrogen. For example, the events of algal blooms emit volatile sulphur, which has an impact on cloud formation. Furthermore, the accumulation of algal biomass is considered to be an ecological hot-spot that is essential to the food chain as it is consumed by a multitude of animals (Schleyer and Vardi, 2020).

### 1.1.2 Causes

The rapid and uncontrollable growth of these algal blooms, whatever their category, can be attributed to several factors, which are categorised into two distinct groups: human causes and natural causes (Table 1.1).

Table 1.1: Summary of the main causes of algal blooms (Zohdi and Abbaspour, 2019).

| Natural | Human |
|---|---|
| Changes in water temperature | Accumulated foods in aquaculture pools |
| Current systems | Discharge of ships' ballast water |
| Dust storms | Discharge of ships' sewage |
| Earth activities | Eutrophication |
| Large climatic oscillation | Oil spills |
| Light | |
| Low water levels | |
| Ocean eddies | |
| Rainfall | |
| Salinity | |
| Turbulence | |
| Upweilling | |
| Vitamins | |
| Water surface temperature | |
| Wind | |

Human causes are associated with specific economic activities, primarily stemming from the influx of industrial, agricultural or urban waste. This waste is rich in nutrients that are useful for algae growth, such as nitrogen (N) and phosphorus (P) compounds ($NO_3$, $PO_4$, and $NH_4$ being the most available forms; Heller *et al.*, n.d.). These are found along coastlines thanks to wind transport or directly via rivers. In fact, the concentration of phosphorus in coastal waters is 15 times higher than normal. In addition, these high nutrient concentrations have an impact on all phytoplankton species, both non-toxic and toxic (Zohdi and Abbaspour, 2019). However, these can be limiting at a certain point. In fact, the optimum N:P ratio for the best growth is 9:1. Beyond this, one of the two elements becomes limiting for the development of the algae (Heller *et al.*, n.d.). Apart from the release of waste, other causes are linked to human activity, such as aquaculture and problems linked to maritime transport. Ballast water and waste water can carry phytoplankton species or nutrients that can increase the risk of algal blooms. Finally, oil spills caused by oil tankers or other types of vessel, as well as those caused by offshore extraction stations, can contribute to the growth of certain species such as the *Noctiluca dinofagellate* (Zohdi and Abbaspour, 2019).

On the flip side, natural causes are more numerous and diverse, falling into three main categories. The first category pertains to the physico-chemical state of the water, with key factors including water surface temperature, fluctuations in water temperature, and salinity. The second category encompasses causes related to nutrient inputs, akin to human-induced factors, involving elements such as dust storms, rainfall, wind, and marine currents (Zohdi and Abbaspour, 2019). This second category is often pointed out as the main reason in most regions facing repetitive events such as the Arabian/Persian Gulf (Al Shehhi *et al.*, 2014). Finally, the third category involves specific factors, such as the absence of algal predators or climatic oscillations. Typically, these causes, whether acting alone or in combination, account for the recurrent occurrence of algal blooms (Zohdi and Abbaspour, 2019).

### 1.1.3 Consequences

Given the proportion of algae blooms that are not dangerous, the consequences are usually minor. The only problematic consequences are the smell caused by decomposing algae and the visual nuisance of greenish water. In addition, during such episodes, the abundance of algae on the surface can obstruct the light, causing a lack of light for the species below. However, HABs have many harmful consequences (Zohdi and Abbaspour, 2019). They cause significant damage to the marine environment and the local economy. More specifically, it is the toxic HABs that are the most harmful. The toxins produced by HABs impact marine organisms like fish, shellfish, and mammals, and can also affect humans through the food chain or by directly consuming contaminated water (Zohdi and Abbaspour, 2019). Conversely, non-toxic HABs can indirectly lead to the demise of marine species. As these algae proliferate, they deplete nearly all the oxygen in the water, depriving other species of this essential element. Additionally, they can obstruct fish gills or suffocate corals and other aquatic vegetation ("What is a harmful algal bloom?", n.d.). HABs, whether toxic or non-toxic, also inflict harm on the local economy. They impact economic sectors such as aquaculture and tourism, leading to substantial financial losses (Zohdi and Abbaspour, 2019). Finally, algal blooms in general also have an impact on desalination plants. In addition to the production of toxins by certain species of algae, other types of algae can produce suspended solids and dissolved organic matter in large quantities, posing technical problems for desalination processes (Anderson *et al.*, 2017).

### 1.1.4 Development

The development of algae is similar to that of bacteria. It is divided into four distinct phases. Initially, there is a latent phase where no growth is observed. This is followed by a phase of exponential growth when sufficient nutrients are available. Then, when nutrients become limiting, comes the stationary phase, which consists of a perfect equilibrium. Finally, the last phase is that of cell death when all the cells die little by little (Zwietering *et al.*, 1990; Anderson *et al.*, 2017). As far as toxin production for the HABs category is concerned, it is present in each of the phases. However, depending on the species, this production will vary according to the phase. For example, for a given species, peak production may occur during the exponential growth phase when nutrients are not limiting. For other species, this production peak may occur during the stationary phase (Anderson *et al.*, 2017).

### 1.1.5 Dynamic

The dynamics of algal blooms are complex. In fact, the movements induced also allow the supply of nutrients to the algae and are therefore associated with the possible causes of the formation of algal blooms presented in the Table 1.1 earlier. However, these dynamics are almost identical for all species and allow us to understand the observations made.

Firstly, given the non-mobility of most algae, their movements are regulated by marine currents such as the tides and eddies, as well as the wind. Moreover, even mobile species are dispersed by these principles because they are not strong enough to counter the forces of the currents. These currents can induce a movement such that an algal bloom at point A may have been initiated at point B several thousand kilometres away. In addition, close to the coast, other processes involved in the movement of algal blooms are upwelling and downwelling, which are regulated by the wind. Downwelling brings cold, deep, nutrient-rich water onshore, while upwelling forces surface water towards offshore areas. In addition to the movement induced for algae, the boundary between these two water masses presents a high level of biological activity, which is beneficial to algal blooms (Anderson *et al.*, 2017).

A second important element in understanding the movements of algal blooms is that of fronts (Figure 1.2). During movements forced by marine or wind currents and by other physical and physiological factors, algae tend to accumulate in a zone known as convergence front. This front will be clearly visible on the surface of the water. In fact, thanks to the high concentration of algae and therefore pigments, this front will be the most colourful and visible zone. Furthermore, due to the horizontal and vertical movements, this front will be followed by a plume of less and less colour. This is due to the decreasing concentration of algae and the increase in depth, as the plume is below the surface of the water (Anderson, 2001; Anderson *et al.*, 2017).

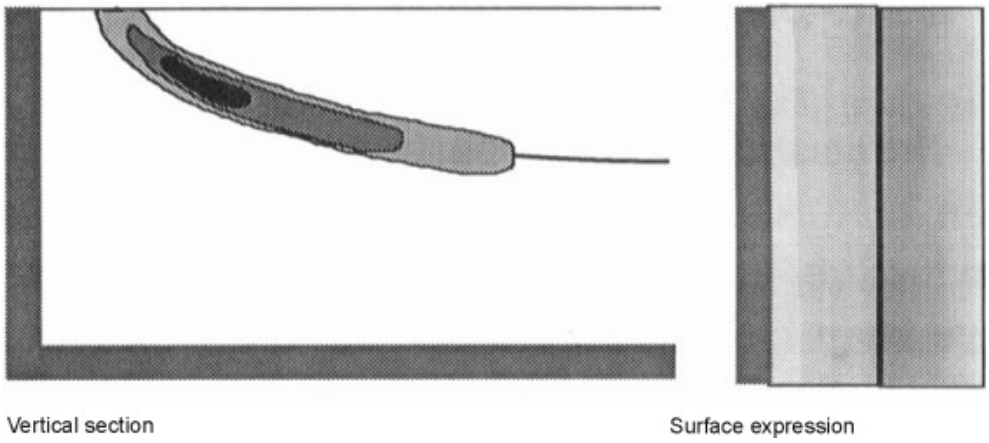



Figure 1.2: Representation of an algal bloom front (Anderson, 2001).

### 1.1.6 Historical events

Because of their less harmful properties, harmless algal blooms are less listed than harmful ones. Nevertheless, some organisations still refer to them when they are easily observable. One such organisation is the Copernicus programme with its 'Image of the Day' articles ("Image of the day | Copernicus", n.d.). However, the articles on algal blooms only cover a small proportion of the actual number of algal blooms. For their part and as mentioned above, HABs are more closely monitored. There are a good number of articles about them and databases listing them. Moreover, global warming and all the changes it entails could contribute to an increase in the number of algal blooms and therefore, by definition, HABs. However, while some links have been clearly established, others are less clear-cut. Indeed, it was only in 2019 that the Intergovernmental Panel on Climate Change recognised a direct link between climate change and HABs (Gobler, 2020). Databases have therefore become a practical tool for observing and understanding their evolution more easily. These include the UNESCO database listing the number of HABs observed in 12 regions (Figure 1.3). However, as not all HABs are recorded, the figures available here should be considered as a low average. Despite this, it is possible to identify certain recurring and important events (Hallegraeff, 2021).

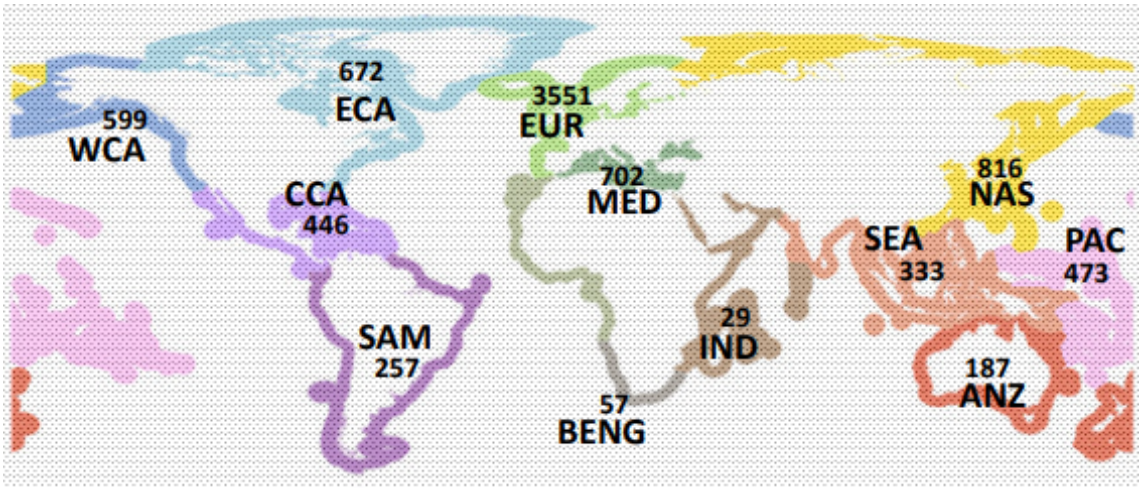



Figure 1.3: Number of recorded HABs for 12 distinct geographical regions. (ANZ = Australia New Zeland. BENG = Benguela Current. CCA = Central America Caribbean. ECA = East Coast America. EUR = Europe. IND = Indian Ocean. MED = Mediterranean. NAS = North Asia. PAC = Pacific. SAM = South America. SEA = South East Asia. WCA = West Coast America) (Hallegraeff, 2021).

In the category of HABs occurring annually, notable instances take place on the West Florida Shelf. The responsible algae in these occurrences has been identified as *Karenia brevis* (Hu *et al.*, 2005). This unicellular microalga belongs to the *dinoflagellate* family and is recognised for its production of brevetoxins, neurotoxins that affect the nervous system of aquatic animals and the respiratory system of humans ("Red Tide", 2018). These blooms seem to originate offshore and are transported to the shore by winds and currents, where they thrive in nutrient-rich waters. The resulting impact on the area is substantial, affecting various economic sectors and impeding the region's overall economic growth (Hu *et al.*, 2005).

An instance of a prolonged HAB is the one that impacted the Gulf of Oman and the Arabian/Persian Gulf region. It emerged at the end of August 2008 off the coast of the United Arab Emirates, extending into the Gulf through the Strait of Hormuz and persisting until August 2009, lasting for a year in the region. The primary species identified in this event was *Cochlodinium polykriloides*. This marked the first occurrence of this species in the region. Similar to the previously mentioned example, the HAB originated offshore and then migrated closer to the shore through sea currents, where it further developed and spread. The prolonged duration of this event can be attributed to various factors, including upwelling and runoff from the few rivers in the region. Due to its extensive duration and magnitude, this event affected over 1,200 kilometers of coastline, resulting in the death of thousands of tons of marine species and causing additional damage, including the shutdown of some desalination plants (Zhao and Ghedira, 2014).

## 1.2 Detection of algal blooms

It therefore seems essential to be able to detect and track down algal blooms. Several types of method exist for this purpose. It is possible to detect the presence of algal species through field campaigns in which samples are taken and analysed. However, this technique is limited for rapid detection because of the time required for sampling and processing (Sellner *et al.*, 2003). Another way is to use optical detection instruments (Cullen *et al.*, 1997) but also optical remote sensing (Ahn *et al.*, 2006; Cao *et al.*, 2021) In addition, it is possible to combine several techniques to increase detection, such as combining satellite images with a digital model (Zhao and Ghedira, 2014). Among all the possibilities, one method stands out: optical imagery, which is currently used in most algal bloom detection methods. This is due to the size and colour of algal blooms, which are easily distinguishable from water and therefore visible from space (Anderson *et al.*, 2017). Moreover, this allows faster results and real-time detection, which is not possible with area sampling (Richardson, 1996).

### 1.2.1 Remote sensing

Remote sensing is a powerful method applicable to various analyses, involving the acquisition of information about the object under observation, in this context, the Earth. Satellite images are commonly employed for this purpose, targeting either terrestrial landscapes or marine environments. Furthermore, two distinct types of sensors are used: radar sensors and optical sensors, each serving distinct purposes and offering unique advantages. The choice between radar and optical sensors depends on the specific objectives of the analysis. For the detection of algal blooms, optical satellite products are the preferred choice.

Optical sensors operate using shorter wavelengths compared to radar sensors, encompassing the visible, near-infrared, and mid-infrared spectral bands. These bands range from 400 to 1600 nm. In contrast, radar sensors use spectral bands with wavelengths in the microwaves to decimeter range (Figure 1.4). Numerous optical sensor satellites are in operation, each designed for specific missions and characterised by various factors, including resolution and revisit period. Resolution varies significantly among different sensors based on their intended applications. Optical sensors rely on the reflection of light from the ground to construct their images. Consequently, images captured by optical satellites typically resemble what is visible to the human eye, with the exception of near- and mid-infrared images.

Near-infrared observations provide insights into the chlorophyll activity of various plant types (Steidle Neto *et al.*, 2017), while mid-infrared observations offer information on soil health (Sanderman *et al.*, 2020). However, because these sensors are based on the visible range and its vicinity, they come with certain drawbacks. Adverse weather conditions can render images unusable, and it is impossible to acquire images at night, as they would appear black (Alvarez-Vanhard *et al.*, 2021).

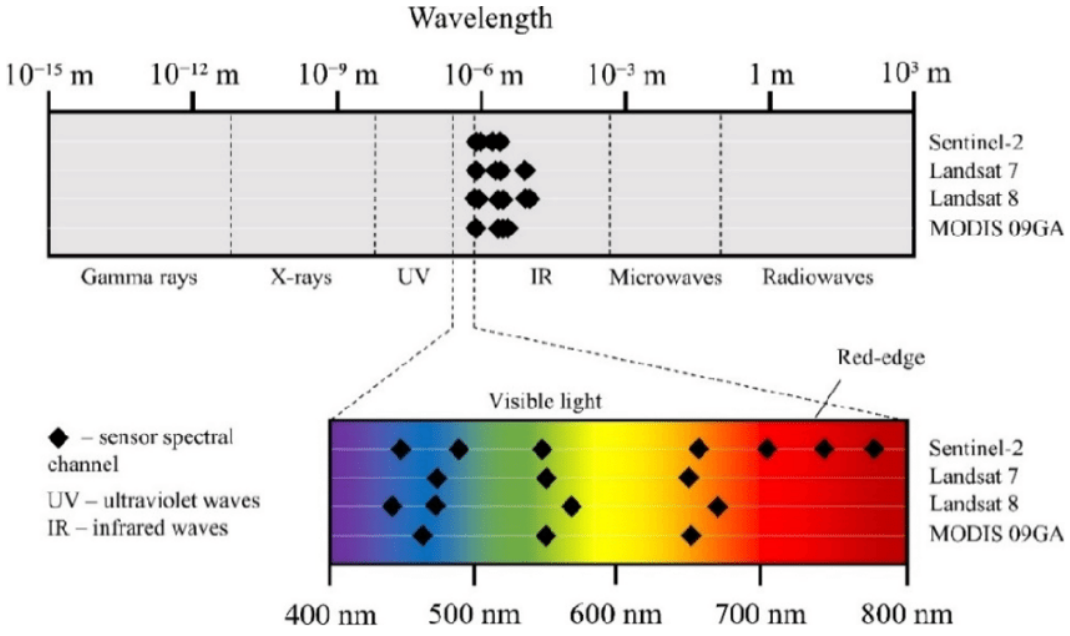



Figure 1.4: Range of spectral bands for certain optical satellites (Radočaj *et al.*, 2020)

### 1.2.2 Chlorophyll

As previously mentioned, near-infrared optical imaging allows the observation of chlorophyll activity (Steidle Neto *et al.*, 2017; Hu, 2009). Short-term variations in this activity can be used to detect abnormal algal bloom events, such as HABs (Wang *et al.*, 2021). However, merely capturing images is insufficient to highlight an algal bloom. Raw data must undergo processing, involving various techniques and algorithms. One notable algorithm for highlighting algal blooms is the Ocean Color (OCx) algorithm, with the latest version being OC6 (O'Reilly and Werdell, 2019). Nevertheless, this method faces a significant limitation in the detection of algal blooms. In regions where observed waters have high concentrations of colored dissolved organic matter, chlorophyll tends to be overestimated (Hu *et al.*, 2005). Therefore, it is crucial to integrate this method with others to enhance and refine the accuracy of the results.

### 1.2.3 ERGB images

Another complementary method involves the use of enhanced RGB images (ERGB). This type of image is built upon the principle of RGB images. An RGB (Red - Green - Blue) image, also known as a true-color image, assigns three values to each pixel-one for blue, one for green, and one for red (Kumar, n.d.). Similarly, an ERGB image operates on the same principle. The distinction lies in the fact that, rather than representing the image in the way we see things, the bands used for the three parameters are chosen to highlight specific elements. Consequently, the image appears with altered colors, but the targeted object stands out distinctly due to its color. In the context of satellite images, the available bands are used to create such products (Zhao and Ghedira, 2014).

### 1.2.4 Detection index

Finally, indexes created from multispectral bands can be used to highlight different types of objects or features. In the context of algal blooms, several indexes can be used, such as the Normalized Difference Vegetation Index (NDVI) or the Floatting Algae Index (FAI) (Cao *et al.*, 2021). These indices have shown their usefulness but also their weaknesses in detecting algal blooms. The NDVI is affected by atmospheric conditions such as clouds and their humidity, but also by turbidity of the water surface (Fan *et al.*, 2020). The FAI, on the other hand, seems to be less sensitive to atmospheric conditions but is still sensitive to turbidity (Hu, 2009; Cao *et al.*, 2021). Finally, the combination of these indices can lead to better results (Xu *et al.*, 2021). However, these combinations have been designed for specific regions (Cao *et al.*, 2021).

CHAPTER

# TWO

# OBJECTIVES

As mentioned in the introduction, algal blooms are emerging as a significant issue due to global warming, and whether harmful or not, they impact the marine environment worldwide. The Arabian/Persian Gulf, in particular, is grappling with the consequences of algal blooms, posing a threat to the operation of desalination plants. To tackle those issues, there is a need for an active and automated detection system to forecast and mitigate the potential disruptions caused by these blooms. This thesis aims to explore the feasibility of implementing a remote sensing system using the open-access satellite image service provided by the European Space Agency (ESA). Several specific objectives for this project are outlined below:

- Create a system based solely on optical satellite imagery from the Sentinel-2 and Sentinel-3 satellites.
- Exploring the use of OTCI via Sentinel-3 to detect days of algal blooms.
- Combine ABDI and NIR from Sentinel-2 to accurately delineate algal blooms.
- Validate the model on the basis of historical events.
- Automate the entire process.
- Apply this process on a daily basis with an early warning if an algal bloom is detected, and send primary information about it.
- Create the embryo of an operational system globally applicable.

CHAPTER

# THREE

# MATERIALS AND METHODS

## 3.1 Zone of interest

The Arabian/Persian Gulf (Figure 3.1), herafter the Gulf, represents an epicontinental semi-enclosed sea within the Indian Ocean. It shares borders with eight countries, serving as a natural boundary between the Arabian Peninsula and Iran and Iraq. Spanning an area of over 250,000 km$^2$, the Gulf is situated between 24°N and 30°N latitude and 48°E and 57°E longitude (Vaughan *et al.*, 2019).

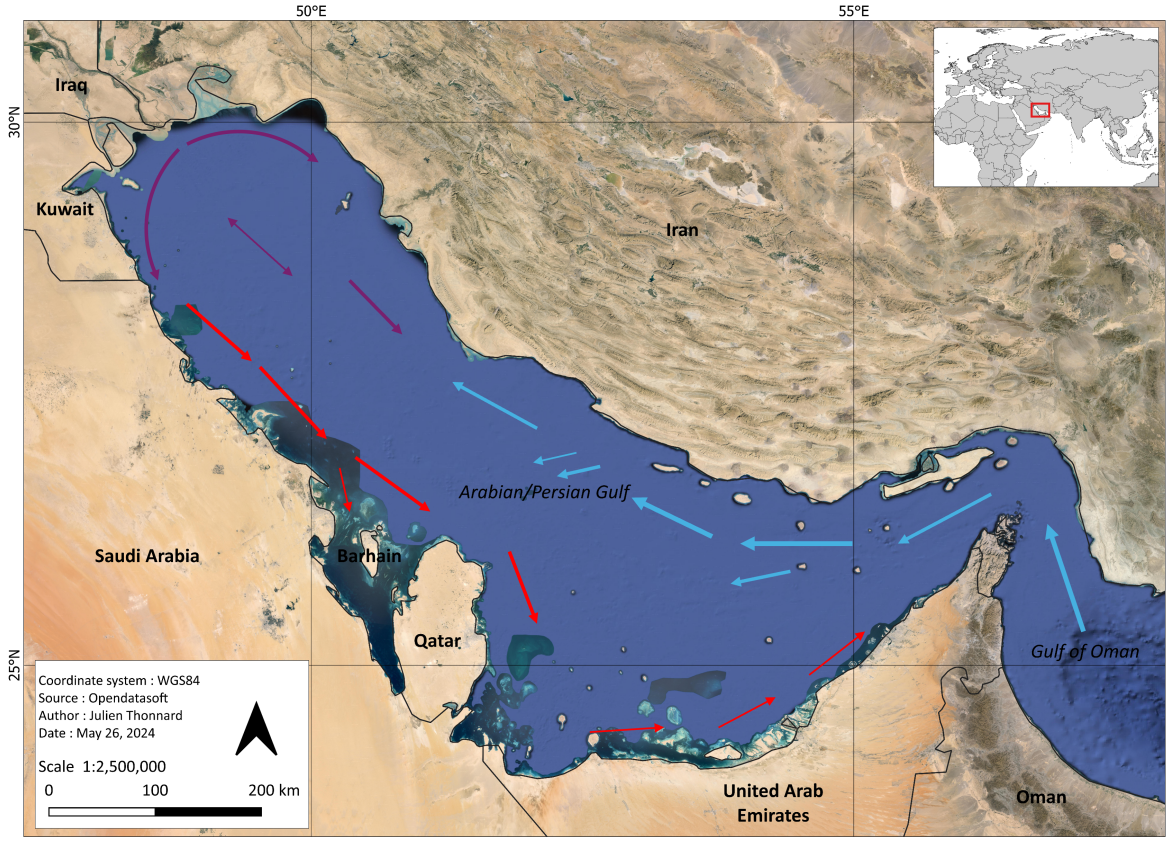



Figure 3.1: Geography of the Gulf with the representation of different marine surface currents in the Gulf, with blue arrows for low-salinity flows, red arrows for high-salinity flows and purple arrows for transition zones (from Michael Reynolds, 1993; Vaughan *et al.*, 2019).

Stretching approximately 940 km in length and narrowing down to 55 km at its thinnest point, it widens to nearly 340 km at its broadest. Notably, the Gulf maintains a relatively shallow nature, with an average depth of 60 meters along the Iranian coast and 35 meters along the northeast coast of the Arabian Peninsula. The Gulf depth increases, to around 100 meters and more, around the Strait of Hormuz, which links it to the Gulf of Oman and the Indian Ocean (Vaughan *et al.*, 2019).

### 3.1.1 Climate

Located within a subtropical high-pressure region (Vaughan *et al.*, 2019), the Gulf is recognised as an arid geographical zone. The climate in this area is marked by two distinct seasons, separated by two transitional periods (Al Senafi and Anis, 2015). During the summer, this region experiences a tropical climate (Vaughan *et al.*, 2019). This season is distinguished by the coexistence of two pressure systems in the Gulf: one characterised by high pressure and the other by low pressure. The pressure gradient generated by these systems results in a robust north-westerly wind (Al Senafi and Anis, 2015) with speeds ranging from 7 to 13 $ms^{-1}$ (Bartlett, 2004), commonly referred to as the summer Shamal (Rao *et al.*, 2003). The summer Shamal brings dust and hot, dry winds (>51°C), known as Simoom, meaning poison in Arabic. In general, this season is marked by an average temperature of 41°C, relatively low humidity at 21%, and no precipitation (Al Senafi and Anis, 2015). The winter season is likewise notable for robust winds known as the winter Shamals. These winds are instigated by the existence of a quasi-stationary Siberian low-pressure system, which imparts cold and powerful winds (Al Senafi and Anis, 2015) reaching speeds of up to 20 m/s across the central surface of the Gulf (Thoppil and Hogan, 2010). Winter is characterised by lower temperatures (<15°C), a relative humidity exceeding 55%, and a higher frequency of precipitation (Al Senafi and Anis, 2015).

### 3.1.2 Hydrographic situation

The Gulf is categorised as an extreme marine environment partially attributable to seasonality and the other characteristics outlined above (Al Senafi and Anis, 2015). The sea surface temperature exhibits a considerable range, fluctuating from 12°C in winter to exceeding 36°C in summer (Coles, 2003), resulting in a notably high annual average. The combination of elevated summer temperatures and the shallow basin depth strongly promotes evaporation in the region. Notably, major rivers in the area, such as the Tigris, Euphrates, and Karun, converging in the north at Shatt Al-Arab, have relatively modest water flow (Al Senafi and Anis, 2015). Despite the contribution of smaller rivers, the collective freshwater input amounts to 110 $km^3$ per year (Khan *et al.*, 2002). However, the estimated annual evaporation from the basin ranges between 360 and 1250 $km^3$ (Michael Reynolds, 1993). These dynamics result in a negative mass balance, leading to an increase in the basin's salinity (Al Senafi and Anis, 2015), with an average value of 42 ppt and a peak exceeding 50 ppt in the southern part (John *et al.*, 1990).

To conclude, an examination of surface currents in the Gulf is essential (Figure 3.1). The primary inflow occurs at the Strait of Hormuz. Inside, the surface low-salinity circulation follows the Iranian coast in a north-western direction (Vaughan *et al.*, 2019). This current is intensified during summer and weakened in winter due to the influence of the Shamal. The Shamal, fostering north-westerly circulation, contributes to the formation of a southerly surface current along the Shatt Al-Arab coast. The wind patterns in the area ultimately induce a northerly return flow in the northern central region of the Gulf (indicated by the double purple arrow). This return flow counteracts the surface water flow, resulting in a low-energy zone characterised by minimal net drift and sensitivity to winds. Along the Arabian Peninsula's coast, a southerly flow descends along the Gulf. This flow stagnates east of Qatar, where evaporation rates are high. Eventually, the flow densifies towards the Strait of Hormuz (Michael Reynolds, 1993).

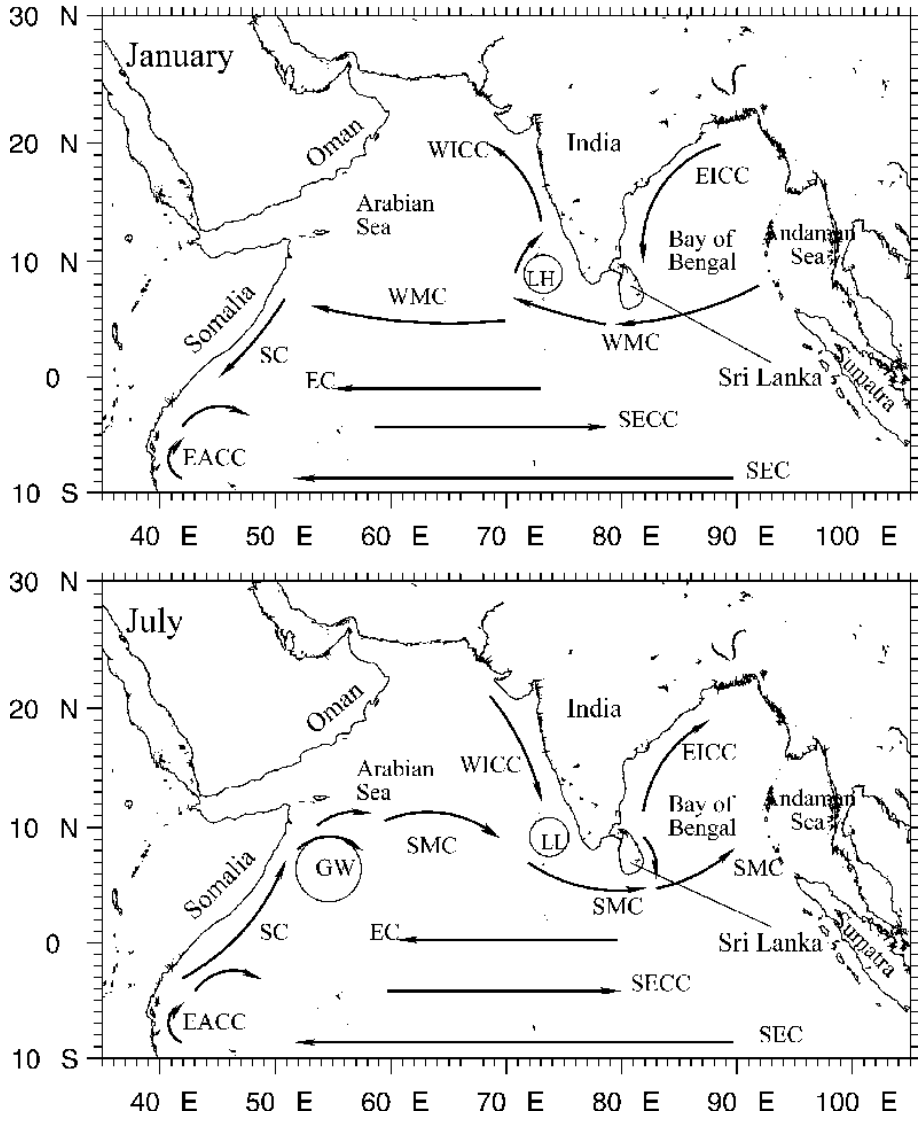



Figure 3.2: Representation of summer and winter marine surface currents in the Indian Ocean (Shankar *et al.*, 2002).

Nevertheless, it is crucial to consider these currents within the broader context of the Indian Ocean. Since the Gulf is an epicontinental semi-enclosed sea of the Indian Ocean, the regional Indian Ocean circulation play a significant role in shaping the Gulf's current system. The Indian Ocean system is intricate and comprises multiple currents (Figure 3.2). However, the focus here is on the northern part of the ocean, given the location of the Gulf. The direction of surface currents varies seasonally, with eastward flows in summer and westward flows in winter, commonly referred to as monsoon currents. This seasonal dynamic is particularly relevant in the context of algal blooms. Figure 1.3 illustrates that few algal blooms (harmful in the case of the image) are detected in the Gulf zone, whereas the adjacent Southeast Asian zone frequently experiences such events. This distinction, coupled with the seasonality of surface currents in the northern Indian Ocean, suggests a potential influence on the occurrence of these events. It is conceivable that an algal bloom originating on the east coast of India could reach the Gulf due to winter marine currents in the Indian Ocean (Shankar *et al.*, 2002).

### 3.1.3 Current geopolitical situation

The Gulf has garnered significant attention as a geographical area for several decades. Beyond the numerous armed conflicts contributing to the current instability in the region (Schmid, 2020), the global focus on the oil market is particularly noteworthy. In 2020, the members of the Gulf Cooperation Council alone accounted for one-third of the world's oil production ("Oil production", n.d.). Given its strategic location, sea transport stands out as the primary method for oil transportation. However, this mode of transport gives rise to various maritime environmental concerns, which are further exacerbated by the geopolitical tensions in the Middle East and the recent attacks by Houthis militias on maritime traffic in the Red Sea.

In an effort to enhance environmental oversight, the Gulf States established the Regional Organisation for the Protection of the Marine Environment (ROPME). ROPME's primary objective is to safeguard the shared maritime areas from marine pollution and external pressures that could impact the region's activities. To achieve this goal, and with the support of robust partnerships such as UNESCO and the World Health Organisation, the organisation has intensified its efforts. These initiatives encompass a range of activities, including remote sensing (Environment, 2017).

In the context of marine pollution and heightened environmental concerns, the automatic detection of algal blooms, whether harmful or benign, seems likely to play a crucial role in monitoring the Gulf's ecosystem.

## 3.2 Algal blooms detection

An algal bloom can be detected thanks to some of these fundamental characteristics. Firstly, the main characteristic that enables detection is the fact that it will tend to be found at the ocean surface. As mentioned in Chapter 1, due to the various sea movements and the influence of the wind, the concentration of algae will be higher at the surface. Moreover, these movements induce high concentrations in areas known as convergent fronts. Secondly, because the color of algal pigments differs from that of water, fronts will appear brightly colored, depending on the algal species. Finally, their spatial and temporal size makes them easy to observe from satellites.

The model that has been developed attempts to make the most of these characteristics to detect any algal blooms that may appear in the Gulf. It is a multi-stage model with various stages briefly described here and further elaborated in the following sections.

### 3.2.1 Satellites imagery

The model uses satellite imagery from the Sentinel-2 and Sentinel-3 constellations. Each of these constellations consists of two satellites: Sentinel-2A and 2B ("Sentinel-2 - Satellite Description - Sentinel Online", n.d.) and Sentinel-3A and 3B ("Sentinel-3 - Mission Summary - Sentinel Online", n.d.).

The European Space Agency (ESA) provides two levels of data for Sentinel-2. The first level, Level-1, is further categorised into two sub-levels: Level-1B and Level-1C. Level-1B offers radiance data of the upper atmospheric layer in sensor geometry, while Level-1C provides the reflection at the top of the atmosphere in map geometry. This correction makes it possible to switch from images that do not contain location information that cannot be used by mapping software to usable corrected images. The second level, Level-2, is divided into three sub-layers. Level-2A provides corrected surface reflections of the atmosphere in map geometry. Although Level-2H and 2F exist as other sub-levels, they are not pertinent to the scope of this work ("Sentinel-2 - Data Products - Sentinel Handbook", n.d.). In general, Sentinel-2 is widely used because of its very fine resolution, ranging from 10 m to 60 m ("S2 Mission", n.d.).

In the case of Sentinel-3, the model uses exclusively the Ocean and Land Colour Instrument (OLCI) sensor in Full Resolution (FR). Similar to Sentinel-2, two main levels are available for public use. Level-1 delivers radiometrically-corrected, calibrated, and spectrally-characterised measurements at the top of the atmosphere. Meanwhile, Level-2 provides products directly derived from Level-1 data, focusing on either submerged land or water. On the other hand, Level-1 bands are expressed in radiance, whereas Level-2 bands are presented in reflectance ("Sentinel-3 - Data Products - Sentinel Online - Sentinel Online", n.d.). Sentinel-3A Level-1 images have been available since mid-2016, while Sentinel-3B Level-1 images have only been available since mid-2018. In addition, a Sentinel-3 image covers more than $1.8 \times 10^6$ km$^2$ made possible by a coarser resolution where one pixel corresponds to an area of 300 m$\times$300 m, or 0.09 km$^2$. This resolution, combined with the presence of two satellites in orbit, means that revisit times at Gulf level are less than 1.4 days for the same area ("S3 Mission", n.d.). This makes Sentinel-3 very interesting for the development of a daily monitoring system for a large area such as the Gulf.

These two optical image sources are not the only ones that can be used. Other possibilities exist, such as SeaWIFS or Landsat 8. However, the latter offer less attractive advantages. SeaWIFS, for its part, has been inactive since 2010 (Earth Science Data Systems, 2024), while Landsat 8 has just 11 bands, compared with 13 for Sentinel-2 ("S2 Mission", n.d.) and 21 for Sentinel-3 ("S3 Mission", n.d.), and a resolution ranging from 15 m to 100 m, compared with 10 m to 60 m for Sentinel-2. Furthermore, the optical bands used for part of the following analysis are between 680 nm and 750 nm, and Landsat 8 has no bands in this wavelength range ("Landsat 8 | Landsat Science", 2021). The combination of the amount of imagery available from Sentinel-3 and the accuracy of Sentinel-2, combined with the band abundance of both, makes them a suitable combination for this analysis. This is why in the context of this project, the selected images will be sourced from Sentinel-2 Level-2A and Sentinel-3 OLCI Level-1 FR datasets. The download process will be facilitated through the online Copernicus browser ("Copernicus Browser", n.d.).

### 3.2.2 Model calibration images

The development of the detection system necessitates a database for testing and validation purposes. The images for this database have been sourced from specific documents in the literature that reference algal blooms observed (see summary in Table 3.1). In addition, eight images with no algal blooms were selected to validate the Sentinel-3 process (see summary in Table A.1). Two instances of these events are visually described in Figure 3.3. Due to the method's independence from meteorological data and the limited availability of papers reporting such events in the Gulf post-2016, the selected events span various locations worldwide.

Table 3.1: Summary of satellite images used for the validation process with the presence of algal blooms.

| Satellite | Location | Date | Source |
|---|---|---|---|
| S2A_L2A T40RDP | Strait of Hormuz & Gulf of Oman | 2019-05-04 | NASA Earth Observatory |
| S2A_L2A T40REP | Strait of Hormuz & Gulf of Oman | 2019-05-04 | NASA Earth Observatory |
| S2A_L2A T40QEG | Arabian Sea | 2024-02-24 | Copernicus - Image of the day |
| S2A_L2A T40QEH | Arabian Sea | 2024-02-24 | Copernicus - Image of the day |
| S2A_L2A T40QFG | Arabian Sea | 2024-02-24 | Copernicus - Image of the day |
| S2A_L2A T40QFH | Arabian Sea | 2024-02-24 | Copernicus - Image of the day |
| S2B_L2A T11TKE | Pyramid Lake, US | 2023-10-08 | Copernicus - Image of the day |
| S3A_L1_EFR | Gulf of Oman | 2017-03-04 | Mcneil, n.d. |
| S3A_L1_EFR | Gulf of Oman | 2017-03-05 | Mcneil, n.d. |
| S3A_L1_EFR | Gulf of Oman | 2017-10-15 | Dolatabadi *et al.*, 2021 |
| S3A_L1_EFR | Gulf of Oman | 2017-10-18 | Dolatabadi *et al.*, 2021 |
| S3A_L1_EFR | Gulf | 2018-05-11 | Duncan, 2018 |
| S3A_L1_EFR | Strait of Hormuz & Gulf of Oman | 2019-05-04 | NASA Earth Observatory |
| S3A_L1_EFR | Strait of Hormuz & Gulf of Oman | 2019-05-05 | NASA Earth Observatory |
| S3A_L1_EFR | Strait of Hormuz & Gulf of Oman | 2019-05-03 | NASA Earth Observatory |

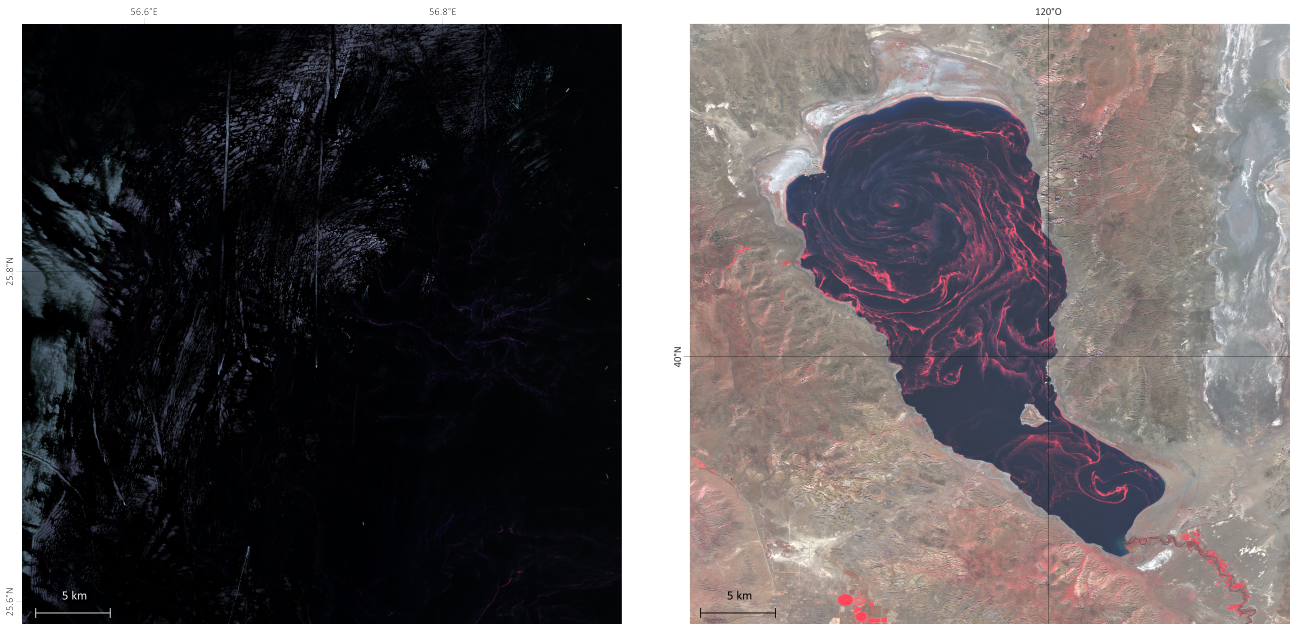



Figure 3.3: Enhanced RGB image based on Sentinel-2 bands 08, 04 and 03 showing algal blooms in the Gulf (left) and in the Pyramid Lake (right).

### 3.2.3 Sentinel-3 initial quick check for blooms

The initial objective is to perform a quick assessment of the area using Sentinel-3 images. If no algal blooms are detected in this step, it allows for a more efficient use of resources because the full process using Sentinel-2 is not necessary. The coverage provided by a Sentinel-3 image facilitates this step. Since mid-2018 and the launch of Sentinel-3B, it has been possible to obtain a daily image covering the area of interest. However, the area covered each day is not the same. It is possible to have the whole Gulf in one image, but it is also possible to have only certain parts of the Gulf for other dates.

Initially, the download will focus on acquiring Sentinel-3 Level-1 FR images. Once the image has been successfully downloaded, it will undergo a series of subsequent processes.

#### Preprocessing

The initial step is preprocessing, which involves shaping the downloaded data according to the required specifications. This step is entirely facilitated by the SeNtinel Application Platform (SNAP) software. SNAP is a tool that allows for image processing and the application of tools specifically designed for Sentinel-1, 2, and 3 ("STEP – Science Toolbox Exploitation Platform", 2024). To expedite this process, two functions, reprojection, and masking of specific areas, have been incorporated into SNAP as a graph.

Reprojection is a crucial step to ensure that any visualisation tool, such as QGIS, can use the image projection system, placing the images accurately. This process involves assigning all layers to the World Geodetic System 84 (WGS84) projection system, which serves as the world's reference system. WGS84 uses the center of the Earth as the origin of coordinates and has an error of less than two centimeters (GISGeography, 2015). The SNAP tool employed for reprojection is the "Reprojection" tool, accessible under the path "Raster -> Geometric -> Reprojection".

Following the completion of the reprojection process, the next step involves masking the image as not all areas of the image are pertinent for the analysis, in other words all pixels that do not correspond to water and pixels that are not usable. To disregard them, the masks provided with the downloaded product are applied to the images. While various masks are available, only those specified in Table 3.2 have been employed in this context.

Table 3.2: Summary of mask layers used

| Full name | Description |
|---|---|
| quality_flags_land | Masks the land |
| quality_flags_fresh_inland_water | Masks lakes and rivers |
| quality_flags_bright | Masks bright pixels |
| quality_flags_invalid | Masks invalid pixels (missing or unusable values) |
| quality_flags_dubious | Mask neighboring pixels of invalids or those coming from a corrupted intrument |
| quality_flags_saturated_Oa10 | Mask saturated pixels for the band 10 |
| quality_flags_saturated_Oa11 | Mask saturated pixels for the band 11 |
| quality_flags_saturated_Oa12 | Mask saturated pixels for the band 12 |

The masks are applied using the "Land/Sea Mask" tool located in "Raster -> Masks -> Land/Sea Mask". As for the last three masks listed in Table 3.2, only these were selected to mask the saturated pixels, as only these bands are required for further processing. Additionally, each of these three masks has been applied to all three bands for ease of use. This approach ensures that a pixel that is saturated and deemed unsuitable for band 10 but not for the other two bands will be disregarded in the analysis. By directly masking all these saturated pixels, the need for prior manipulations is eliminated, allowing for the creation of a single SNAP graph for all three bands rather than three separate graphs. However, since July 25, 2023 for Sentinel-3A and July 18 for Sentinel-3B, only one mask is available for all saturated pixels in all bands. In fact, after these dates, only one mask, not three, needs to be applied to mask saturated pixels in images. The ultimate outcome of this initial step is illustrated in Figure 3.4 where only dark pixels presented on the right are kept for the analysis.

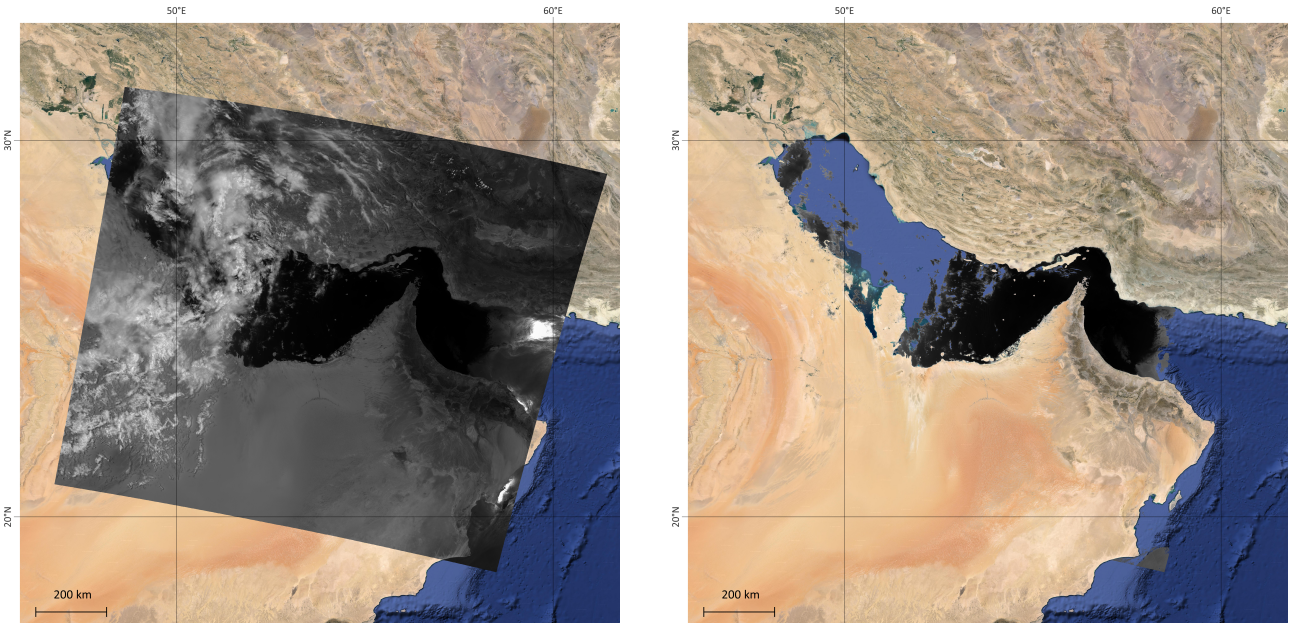



Figure 3.4: Example of an image before (left) and after (right) the application of the masks shown in Table 3.2.

In parallel, SNAP is used to apply a mask along the sides to further minimize the overestimation of chlorophyll implied by the presence of sediment in the water. To this end, the mask layer for the coastline is extracted from each image and expanded by ten pixels, representing three kilometers. The dilated mask is then applied to the image. This additional precaution, combined with the masked areas, ensures that no sediment-loaded zones will impact detection.

**OTCI computation**

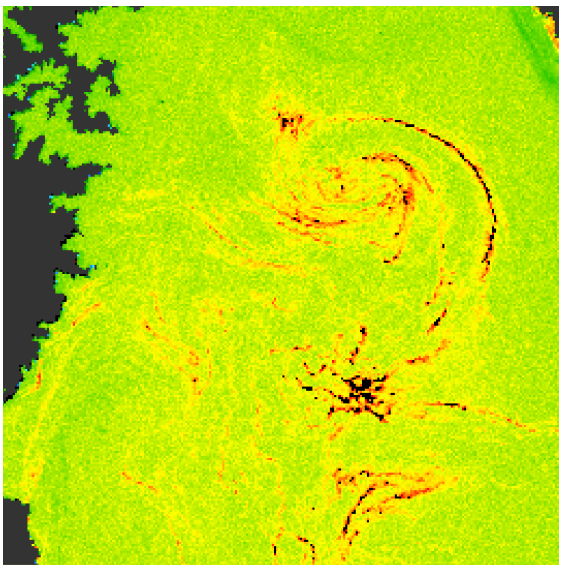
The second phase involves calculating the OLCI Terrestrial Chlorophyll Index (OTCI) given by equation 3.2, which is analogous to the Normalized Difference Vegetation Index (NDVI) given by equation 3.1. However, OTCI is designed to mitigate index saturation issues that may arise when chlorophyll values are exceptionally high. Unlike NDVI-type indices that rapidly saturate ("OLCI Applications", n.d.), OTCI exclusively employs bands in the near-infrared. This approach is advantageous as absorption collapses at longer wavelengths, enhancing the detection of high chlorophyll values ("OLCI Applications", n.d.).

$$NDVI = \frac{R_{842nm} - R_{665nm}}{R_{842nm} + R_{665nm}}, \tag{3.1}$$

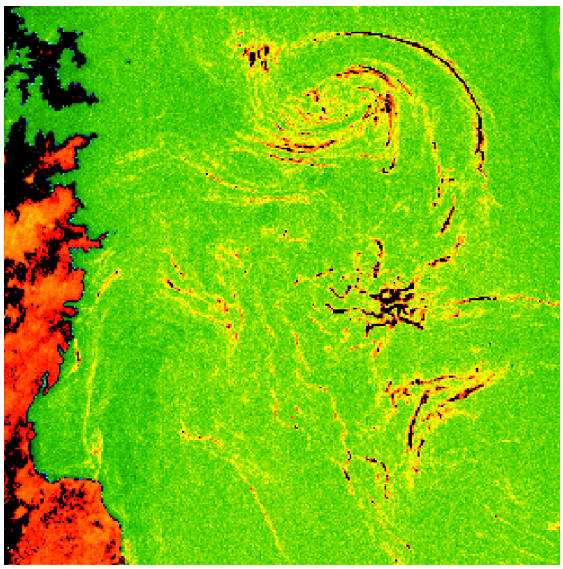
$$OTCI = \frac{R_{753.75nm} - R_{708.75nm}}{R_{708.75nm} - R_{681.25nm}}, \tag{3.2}$$

where $R$ is the reflectance at the indicated wavelength.

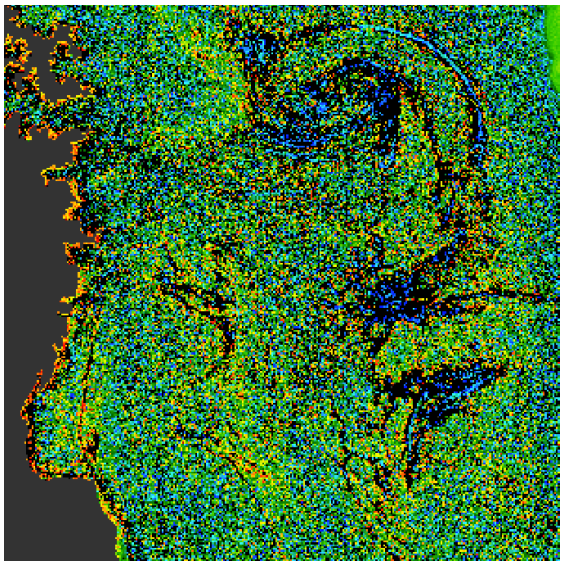
However, as mentioned above, Sentinel-3 Level-1 gives radiance and not reflectance, while OTCI is normally calculated using reflectance. A manipulation is possible via SNAP to switch from radiance, the amount of light observed from an object, to reflectance, the ratio of light received and returned by the object. However, the addition of this step does not enable reprojection and *vice versa*, making the product impossible to read as shown in the first two images (Figure 3.5), where reprojection is applied to a and not b, inducing a shift in the image. Furthermore, as shown in this figure the difference between OTCI calculated via radiance or reflectance seems relatively small. The use of Level-2 images given directly in reflectance therefore seems more appropriate. Nevertheless, calculating OTCI via Level-2 images gives false values. Indeed, pixels within the algal bloom have negative chlorophyll values. Moreover, the delimitation is less precise than using radiance. For all those reasons, OTCI will be computed using the radiance to make processes simpler and faster.



a OTCI computed with radiance.



b OTCI computed with SNAP reflectance.



c OTCI computed with S3 Level-2 reflectance.

Figure 3.5: Comparison between the different way to compute OTCI (blue represents zero chlorophyll, green represents low chlorophyll (water), red represents high chlorophyll, and black represents no chlorophyll).

Ultimately, to show off only the algal bloom, the OTCI scale provided by the Copernicus progam is used. This scale runs from 0 to 5, with the chlorophyll range extending from 1.8 (low chlorophyll content) to 4.5 (high chlorophyll content). For the purposes of this model, a threshold value between 1.8 and 4.5 will be considered, and only pixels exceeding this value will be considered as potentially referring to an algal bloom. However, the chlorophyll present in the water does not necessarily refer to algae. In fact, sediment flows tend to overestimate the chlorophyll content. For this reason, it is necessary to apply a mask to certain areas (Figure 3.6). Indeed, the Gulf faces a fairly high sediment load and some areas, such as the northern part of the Gulf, are constantly exposed to these flows (Alosairi *et al.*, 2024). In addition to these masks, a second threshold value on the number of pixels greater than the first threshold value will be applied to avoid considering a sediment flux as an algal bloom. This double security should make it possible to consider only algal blooms.

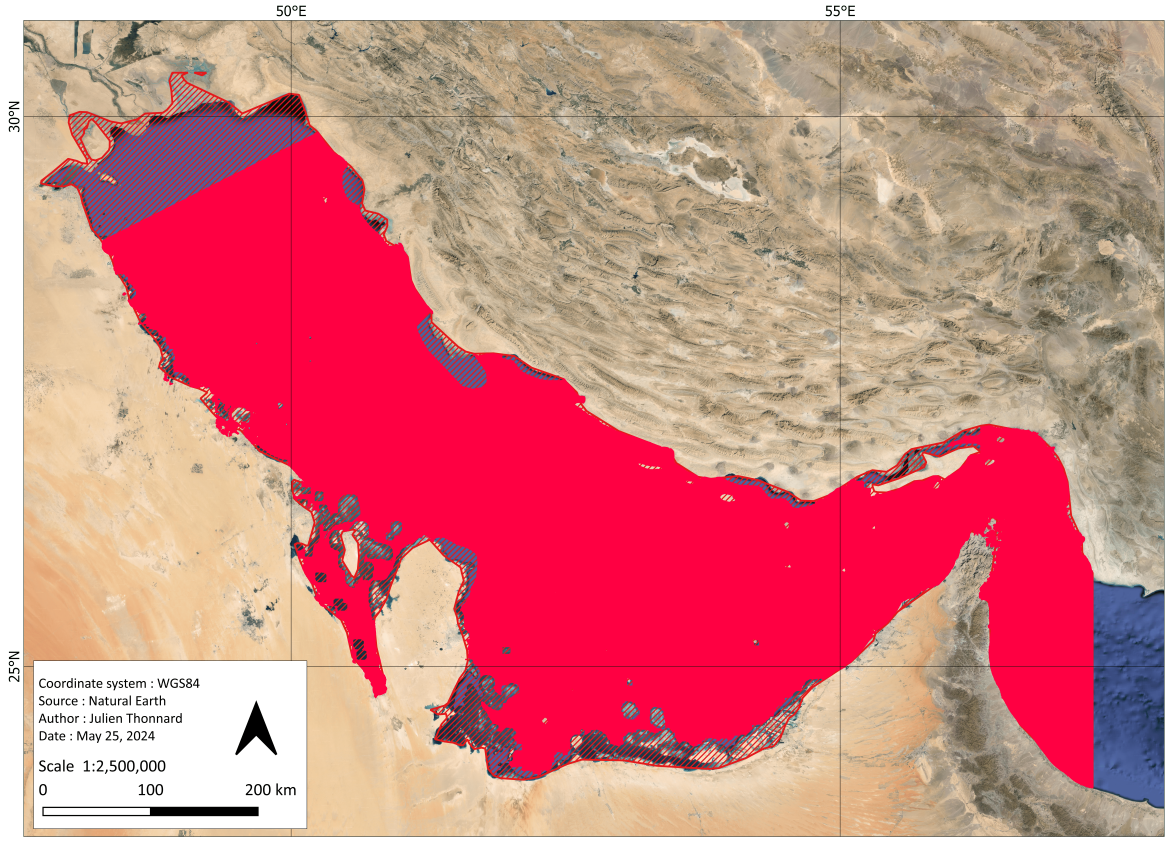



Figure 3.6: Representation of the masked area in hatched red and the area under consideration in red.

Finally, if the day analysed shows an algal bloom, the model will go on to the next step and analyse the Sentinel-2 images. In the opposite case, the process for that date is stopped and the algorithm moves on to the next date.

### 3.2.4 Sentinel-2 process

Once a bloom has been detected on a Sentinel-3 image, the analysis can be refined by using higher resolution Sentinel-2 images.

#### Preprocessing

Before using the images, a series of preprocessing steps is necessary. In contrast to the preprocessing of Sentinel-3 images, the preprocessing of Sentinel-2 images is performed using Python. The steps involved are derived from the course "LBRAT2104 Land Monitoring by Advanced Earth Observation". However, in this case, only steps one and three are applied, which involve (1) resampling the bands from 20-meter to 10-meter resolution and (2) applying the Scene CLassification (SCL) map to mask out unnecessary pixels ("Preprocessing Sentinel-2 images — eo-toolbox 0.1 documentation", n.d.).

The resampling process is split into two parts: first for the SCL map and then for the bands. The SCL map, being categorical data, can only be resampled using the nearest neighbor method. On the other hand, the bands are continuous data, allowing for the application of different resampling methods. In this case, the bilinear method is employed ("Preprocessing Sentinel-2 images — eo-toolbox 0.1 documentation", n.d.). Bilinear interpolation is an interpolation method that uses the four nearest cell centers. These four data points are weighted based on their distance from the unknown point. The average of these weighted values yields the value of the new point, effectively smoothing the data (GISGeography, 2017).

In the second step, the image is masked using the SCL map, which assigns a numerical value to each pixel for each class. The distinct numbers for different classes are detailed in Table 3.3. For this analysis, only pixels corresponding to water and vegetation are retained, while the rest of the pixels are set to NaN and are consequently excluded from further consideration. In addition, in order to consider only the vegetation pixels found in the maritime zones, a second mask for the terrestrial zones is applied, masking the terrestrial vegetation pixels. It is necessary to take vegetation pixels into account because in some cases the chlorophyll concentration of a pixel corresponding to algae is such that it is considered to be vegetation.

Table 3.3: Description of the different classes in the SCL map ("Preprocessing Sentinel-2 images — eo-toolbox 0.1 documentation", n.d.)

| SCL class | Description |
|---|---|
| 0 | No data |
| 1 | Saturated or defective |
| 2 | Dark area pixels |
| 3 | Cloud shadow |
| 4 | Vegetation |
| 5 | Not vegetated |
| 6 | Water |
| 7 | Unclassified |
| 8 | Cloud medium probability |
| 9 | Cloud high propability |
| 10 | Thin cirrus |
| 11 | Snow |

### Bloom detection

The subsequent steps are divided into two parallel processes. The first involves calculating the Algae Blooms Detection Index (ABDI), proposed by Cao *et al.* (2021), which appears promising for algal bloom detection. This index is derived using the following equation:

$$ABDI = \left(R_{red-edge2} - R_{red} - (R_{NIRn} - R_{red})\frac{(\lambda_{red-edge2} - \lambda_{red})}{(\lambda_{NIRn} - \lambda_{red})}\right) - \left(R_{red} - \frac{R_{green}}{2}\right), \quad (3.3)$$

where $R$ gives the reflectance and $\lambda$ is the central wavelength in both cases for the bands mentioned.

As shown in the equation above, ABDI is divided into two components. The first part extracts the algal bloom from the image, and this part alone is effective when the object is present in clear water. However, in the case of turbid water, using the first part alone might not adequately highlight the algal bloom. This is where the second part of the equation becomes crucial. It differentiates algal blooms from turbid water by subtracting the difference between the reflectance for the red band and half the green band. This reduces the index's sensitivity to turbid water, but does not entirely solve the problem. (Cao *et al.*, 2021).

Simultaneously, band number 08 is extracted from the image file. This band corresponds to a wavelength of 842 nm and is one of the bands that captures infrared information. Apart from the initial preprocessing, no further alterations are made to the image. Once these two by-products are obtained, they undergo a post-processing steps. Both layers are standardized, and only values above the threshold value set for each band are retained.

**Layers intersection**

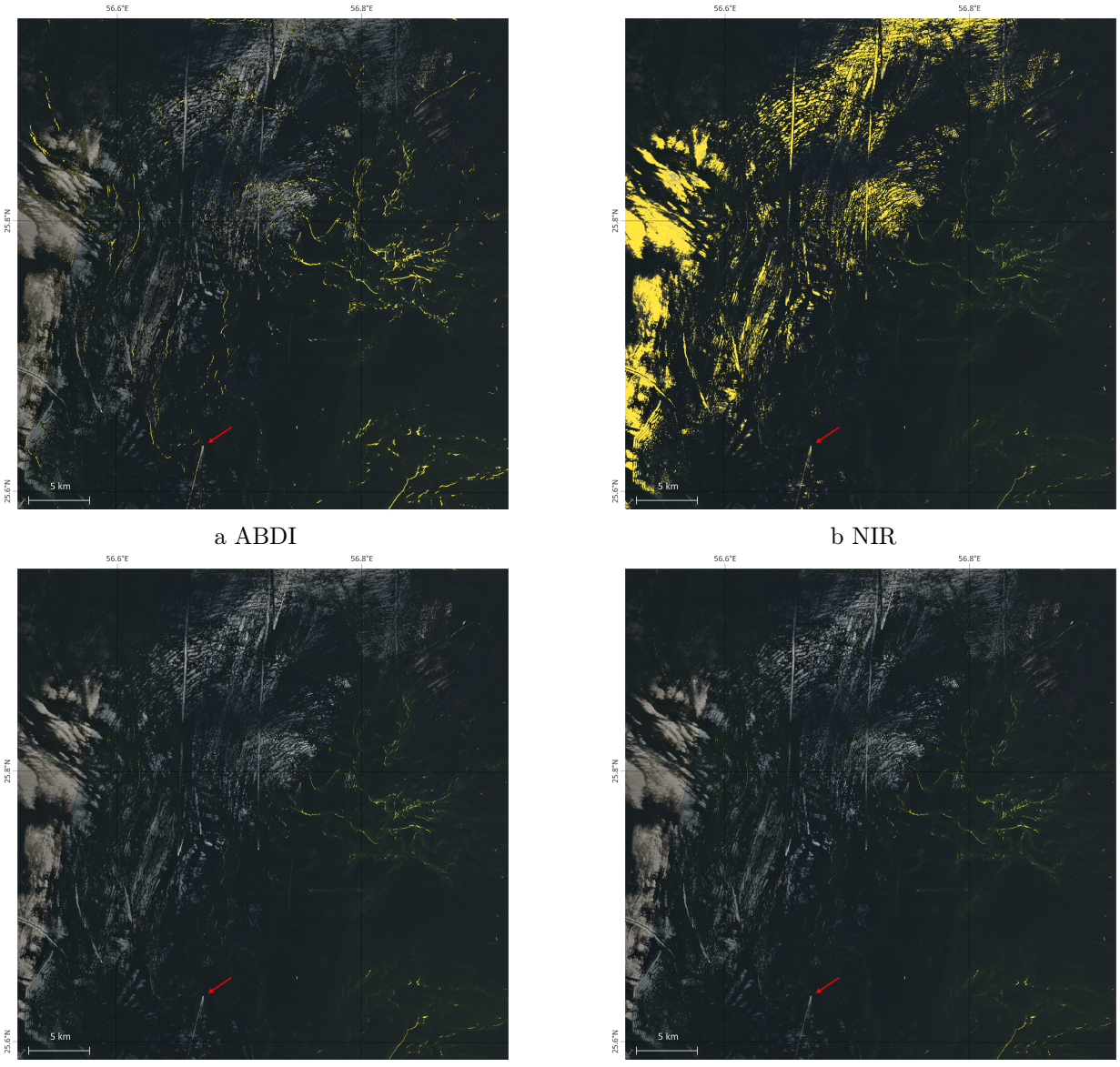
These two analyses highlight potential algal blooms. However, other objects are also considered, such as boats and their wakes, highlight with arrow in the bottom middle of the images in Figure 3.7. The analysis is conducted using optical images which are degraded in the presence of clouds. This underscores the importance of masks. However, masks are not always entirely effective, as the fundamental identification of pixels may be incorrect. For instance, a pixel indicating a cloud might be misidentified as a water pixel. In the analysis, if such misclassifications occur, it can lead to potential errors when calculating ABDI as well as highlighting by near-infrared.

It is then necessary to combine ABDI with a second product, NIR. NIR can detect chlorophyll, but also other elements such as clouds (Chylek *et al.*, 2006). This explains the high number of pixels retained for the analysis of an image over the western part of Gulf of Oman, as indicated in Table 3.4. However, to avoid considering the entire layer for both products, a selection is applied to the layers. They are normalized to mean and standard deviation. Then, a threshold value corresponding to a number of times the standard deviation is established, and only values greater than or equal to this value are retained for further processing.

Then, to selectively highlight only objects corresponding to algal blooms, the intersection of the two previous products is employed. This allows the exclusion of wrong objects considered by ABDI but not by NIR, and vice versa resulting from a reduction in the number of pixels considered (Table 3.4). The resulting layer indicates objects that still have a high probability of being algal blooms. However, some residual pixels may persist in the image. To address this, a smoothing process is applied to the intersection product. Any object consisting of a single pixel is disregarded. This last step allows us to further reduce the areas considered and increase the reliability of areas considered to be algal blooms.

Table 3.4: Comparison of the number and proportion of pixels highlighted at each stage of analysis for the image of 04 may 2019 in the Gulf of Oman.

| Letter | Treatment | Number of highlighted pixels | proportion of highlighted pixels [%] |
|---|---|---|---|
| a | ABDI >1std | 947,002 | 0.785 |
| b | NIR >1std | 5,075,922 | 4.216 |
| c | ABDI intersection NIR | 26,723 | 0.022 |
| d | ABDI intersection NIR + smoothing | 23,065 | 0.019 |




a ABDI

b NIR

c Intersection between ABDI and NIR

d Smoothing

Figure 3.7: Comparison of pixels highlighted (yellow) at each stage of analysis for the image of 04 may 2019 in the Gulf of Oman, with the red arrow indicating the presence of a ship.

Pixel exclusion works well, as shown in Figure 3.7 and Table 3.4, however, pixels that are part of algal blooms are rejected. That is why it is necessary to consider them, and that is what the next step is for.

### Buffer to extend detection

After refining the results, the objective of this step is to expand the mask starting from these few pixels. To encompass the entire algal bloom, a buffer around the pixel is created. Initially, a mask is generated that includes the initial pixels and adjacent pixels. Adjacent pixels are selected iteratively. In the first iteration, only the four pixels adjacent to the initial pixel are considered. In the second iteration, the eight pixels adjacent to the previous four are added, and so on. The number of iterations is defined in the algorithm beforehand. Once the mask is created, it is overlaid on top of the ABDI layer. The two layers are intersected, and only pixels with an ABDI value higher than the threshold value defined upstream of the method and included in the mask are retained. The product of this intersection represents the final output of Sentinel-2 image analysis and can be observed in Figure 3.8.

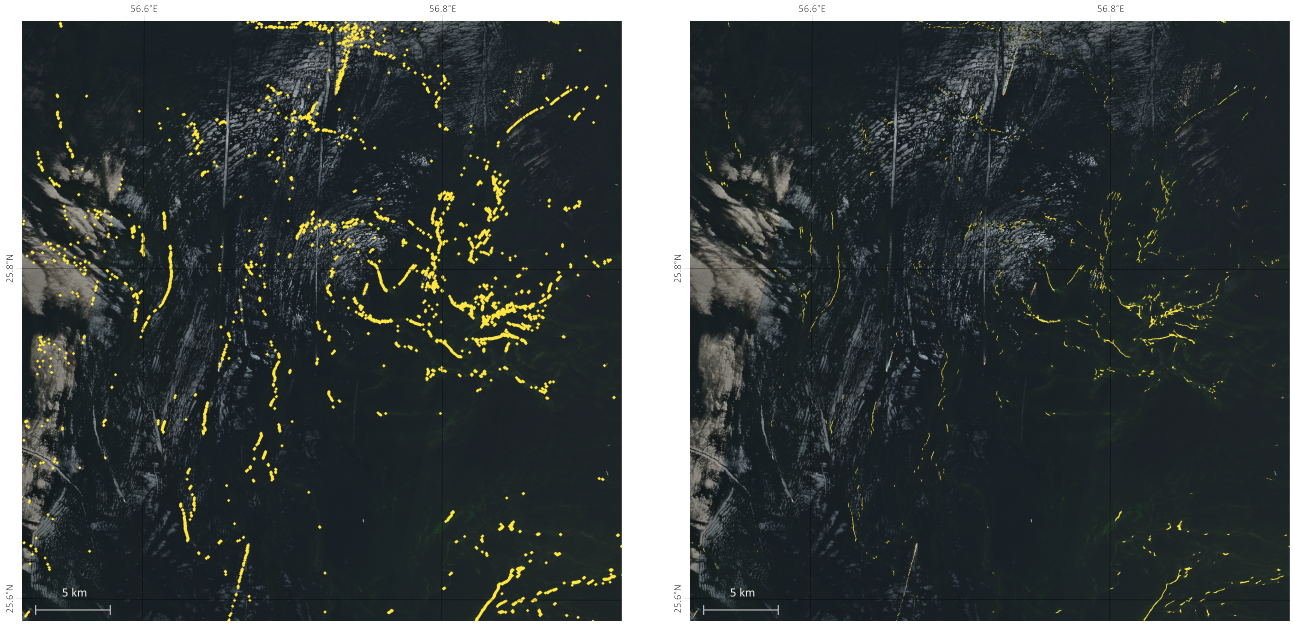



Figure 3.8: Representation of the buffer with 10 iterations (left) and of the result of the intersection between the buffer and the ABDI layer (right) for the image of 04 may 2019 in the Gulf of Oman.

## 3.2.5 Algal bloom visualisation

The ultimate visualization is the outcome of the combined Sentinel-3 and Sentinel-2 analysis. On a date when both satellites provide images, the Sentinel-2 product will be more accurate than the Sentinel-3 one as it has a finer resolution (10 m VS 30 m).

### 3.2.6 Method summary

All the steps described above are shown in the figure below. It summarises the entire method, with links between each step and the various possibilities.

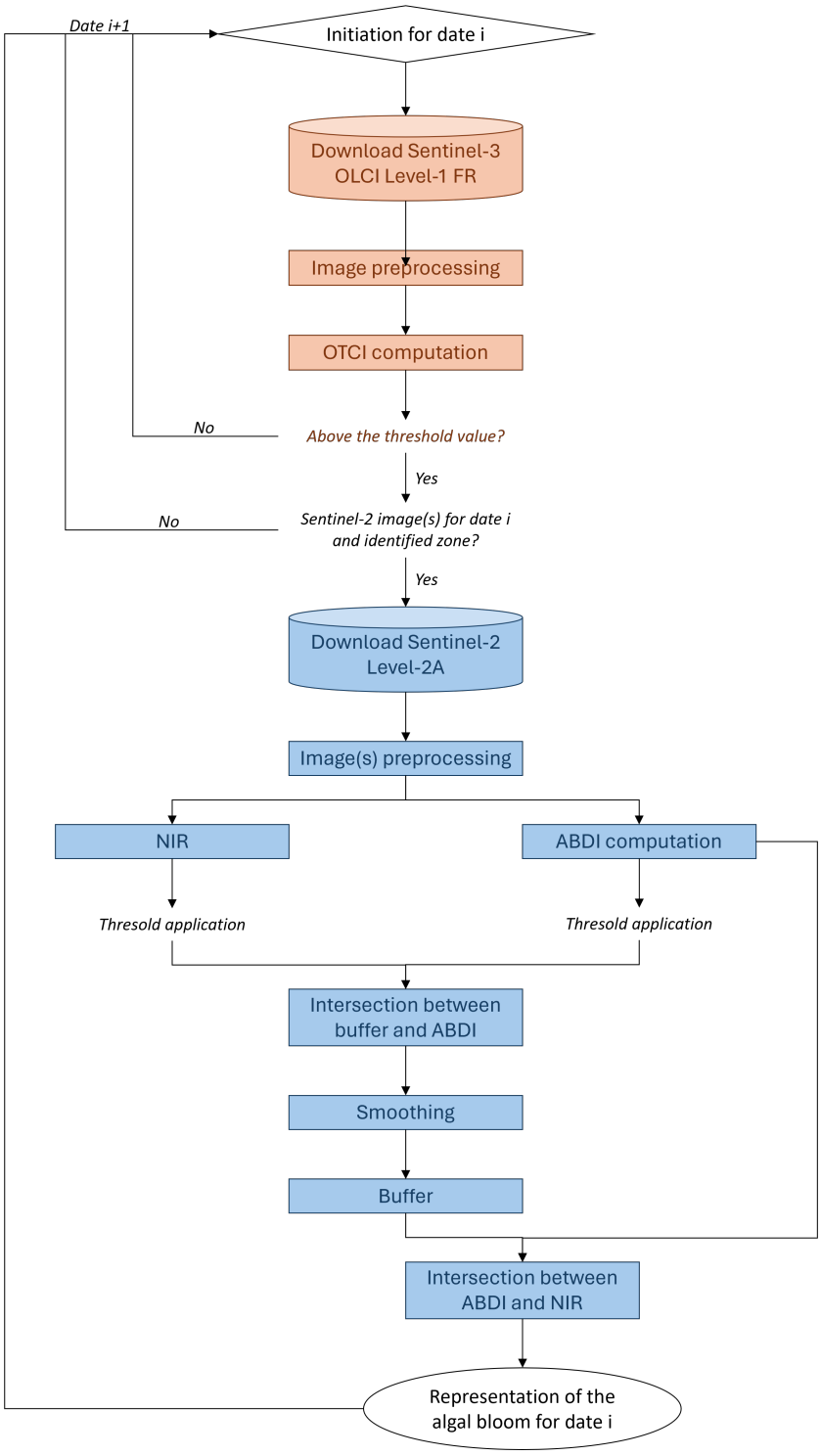



Figure 3.9: General detection scheme with brown boxes corresponding to Sentinel-3 processes and blue ones corresponding to Sentinel-2 processes.

## 3.3 Validation

### 3.3.1 Sentinel-3

The validation of Sentinel-3 results will be binary. To this end, the sixteen Sentinel-3 images presented in Tables 3.1 and A.1 will be used. The aim of this validation will be to see, for a set of two threshold values, whether the eight images where an algal bloom is present are captured, while ignoring the eight images with no algal bloom.

### 3.3.2 Sentinel-2

In the case of this model for Sentinel-2, the validation will involve a performance evaluation using a confusion matrix, which compares the model's predictions with ground truth. The predicted values are presented in the columns, while the actual values are displayed in the rows. The simplest confusion matrix compares two classes with each other (Figure 3.10) and is made up of the following four different cells (Luque *et al.*, 2019):

- True positives (TP) are correctly predicted positive values.
- False positives (FP) are negative real values falsely predicted as positive.
- False negatives (FN) are positive values for the reals that are incorrectly predicted as negative values.
- True negatives (TN) are correctly predicted negative values.

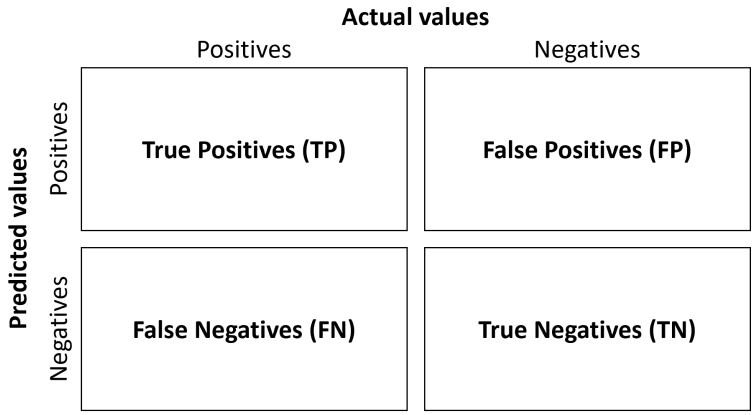



Figure 3.10: Representation of the general appearance of a confusion matrix.

From this confusion matrix, it is possible to derive other metrics to evaluate the model's performance. The first is accuracy, which provides the rate of values correctly predicted by the model. In the case of confusion matrices, accuracy is obtained using the following equation (Luque *et al.*, 2019):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{3.4}$$

However, while this metric provides an initial understanding of the model's performance, it is not sufficient, especially in cases where the database exhibits an imbalance of classes. Therefore, it is necessary to complement the performance analysis with other metrics (Luque *et al.*, 2019).

The second metric, precision, which is defined as the proportion of positive values predicted correctly (Luque *et al.*, 2019):

$$Precision = \frac{TP}{TP + FP}. \tag{3.5}$$

In addition to precision, recall enables the determination of the rate of correctly predicted positive values (Luque *et al.*, 2019):

$$Recall = \frac{TP}{TP + FN}. \tag{3.6}$$

Furthermore, the last two metrics can be combined through the F1-score. This metric combines recall and precision to estimate the accuracy of the model by class (Luque *et al.*, 2019):

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}. \tag{3.7}$$

Finally, the last test we will use is the Receiver Operating Characteristic (ROC) curve. This curve highlights the ability of a classifier to correctly predict positive values, with the least error in predicting these values. In addition, a straight line $y = x$ establishes the prediction of a random classifier. It is therefore important to be above this line for the model to be interesting to use. In order to perform this analysis, two metrics are required: the true positive rate (TPR) and the false positive rate (FPR) (Hoo *et al.*, 2017):

$$TPR = \frac{TP}{TP + FN}, \tag{3.8}$$

$$FPR = \frac{FP}{FP + TN}. \tag{3.9}$$

In the context of this thesis, various performance analysis metrics will be applied to the model results. To achieve this, the Sentinel-2 images presented in Section 3.2.2 will be used in two stages. In the first stage, the method described above will be applied, and the resulting dataset will form the prediction dataset. Secondly, 3,000 random points will be placed along the detected objects to facilitate validation. This step will be carried out in QGIS for ease of use. The scores obtained will then be analysed to assess and predict the model's behavior and applicability. It is essential to point out that greater weight will be given to metrics suitable for unbalanced

databases. As algal blooms are minority elements in the observed images, the proportion of pixels corresponding to an algal bloom will be a minority compared to the rest. Consequently, while accuracy will be considered, its significance will be minimised in the decision-making process.

CHAPTER

# FOUR

# RESULTS

The results will be divided into four main sections, with a presentation of the model's calibration and limitations in the first two sections. This will be followed by the results obtained from the simulation carried out for the year 2023 using the results found in the first two parts. Finally, we present the results of a cause-consequence analysis between the occurrence of an oil spill and that of an algal bloom.

## 4.1 Model calibration

As the model is based on two indexes (OTCI and ABDI) and severeal processing steps with different parameters, a model calibration stage is necessary to optimise each model step for Sentinel-3 and Sentinel-2 images. Proper model parameterisation will enable more accurate detection of algal blooms in the Gulf.

### 4.1.1 Sentinel-3

First, the calibration results for the section including Sentinel-3 images are presented. As a reminder, in this part two hyperparameters are present and variable. The first is for the OTCI value, a dimensionless index, and the second for the number of pixels with an OTCI value greater than or equal to the first threshold. For the first threshold, we take value of 1.8, 2.5 and 4, based on the Copernicus Browser scale. For the second threshold, we considered ten uniformally distributed values between 10 and 100. However, for the OTCI value of 1.8, the number of pixels exceeding this threshold is too high. We hence consider 10 uniformally-distributed values between 100 and 1000. This yields 30 different combinations to test on 16 different images, the first eight show algal blooms and the last eight do not. The results of this analysis are shown in Tables 4.1, 4.2 and 4.3. Each table shows for each image whether the threshold under consideration allows the detection of the algal bloom (in white) or not (in dark grey).

In Table 4.1, we detect all eight algal bloom images with thresholds of 100 and 200. However, with a threshold of 100 or 200 for images with no algal bloom captured on October 17 of 2020, it is predicted as having an algal bloom. The same happens for thresholds up to 500. Using an OTCI value of 1.8 therefore leads to possible errors in the model's prediction.

Table 4.1: Validation results for pixels with OTCI value $\geq 1.8$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image.

| Status | Date | OTCI | Threshold | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
| Bloom | 04-03-2017 | 577 | | | | | | | | | | |
| | 05-03-2017 | 401 | | | | | | | | | | |
| | 15-10-2017 | 232 | | | | | | | | | | |
| | 18-10-2017 | 395 | | | | | | | | | | |
| | 11-05-2018 | 337 | | | | | | | | | | |
| | 03-05-2019 | 704 | | | | | | | | | | |
| | 04-05-2019 | 2509 | | | | | | | | | | |
| | 05-05-2019 | 913 | | | | | | | | | | |
| No bloom | 24-02-2017 | 17 | | | | | | | | | | |
| | 19-06-2019 | 29 | | | | | | | | | | |
| | 17-10-2020 | 599 | | | | | | | | | | |
| | 10-12-2020 | 8 | | | | | | | | | | |
| | 27-06-2021 | 68 | | | | | | | | | | |
| | 13-04-2022 | 83 | | | | | | | | | | |
| | 05-08-2022 | 0 | | | | | | | | | | |
| | 25-06-2023 | 1 | | | | | | | | | | |

When using an OTCI value of 2.5, the results show that for images with algal blooms, thresholds between 10 and 60 lead to a good prediction for all images (Table 4.2). For images without algal blooms, the results show that from a threshold of 50, all images are predicted correctly. It is therefore possible to select thresholds 50 and 60 as parameters that can be used for final detection when the minimum OTCI value is set to 2.5.

Table 4.2: Validation results for pixels with OTCI value $\geq 2.5$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image.

| Status | Date | OTCI | Threshold | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| Bloom | 04-03-2017 | 270 | | | | | | | | | | |
| | 05-03-2017 | 155 | | | | | | | | | | |
| | 15-10-2017 | 77 | | | | | | | | | | |
| | 18-10-2017 | 96 | | | | | | | | | | |
| | 11-05-2018 | 67 | | | | | | | | | | |
| | 03-05-2019 | 193 | | | | | | | | | | |
| | 04-05-2019 | 725 | | | | | | | | | | |
| | 05-05-2019 | 284 | | | | | | | | | | |
| No bloom | 24-02-2017 | 6 | | | | | | | | | | |
| | 19-06-2019 | 0 | | | | | | | | | | |
| | 17-10-2020 | 45 | | | | | | | | | | |
| | 10-12-2020 | 5 | | | | | | | | | | |
| | 27-06-2021 | 9 | | | | | | | | | | |
| | 13-04-2022 | 16 | | | | | | | | | | |
| | 05-08-2022 | 0 | | | | | | | | | | |
| | 25-06-2023 | 1 | | | | | | | | | | |

Finally, if the minimum OTCI value taken into account is 4, only a threshold of ten pixels allows a good prediction of the eight images with an algal bloom and the other eight images without. However, for the dates of October 15 and 18, 2017 and May 11, 2018, the number of pixels highlighted is close to this threshold value of ten, which may not be appropriate for other images.

Table 4.3: Validation results for pixels with OTCI value $\geq 4$, with white cells corresponding to cases where an algal bloom is detected on the image, and dark grey cells to cases where no algal bloom is detected on the image.

| Status | Date | OTCI | Threshold | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| Bloom | 04-03-2017 | 63 | | | | | | | | | | |
| | 05-03-2017 | 26 | | | | | | | | | | |
| | 15-10-2017 | 10 | | | | | | | | | | |
| | 18-10-2017 | 14 | | | | | | | | | | |
| | 11-05-2018 | 15 | | | | | | | | | | |
| | 03-05-2019 | 36 | | | | | | | | | | |
| | 04-05-2019 | 109 | | | | | | | | | | |
| | 05-05-2019 | 39 | | | | | | | | | | |
| No bloom | 24-02-2017 | 0 | | | | | | | | | | |
| | 19-06-2019 | 1 | | | | | | | | | | |
| | 17-10-2020 | 1 | | | | | | | | | | |
| | 10-12-2020 | 1 | | | | | | | | | | |
| | 27-06-2021 | 0 | | | | | | | | | | |
| | 13-04-2022 | 4 | | | | | | | | | | |
| | 05-08-2022 | 0 | | | | | | | | | | |
| | 25-06-2023 | 0 | | | | | | | | | | |

### 4.1.2 Sentinel-2

Secondly, the influence of the method's four hyperparameters on Sentinel-2 images was analysed with several metrics. As a reminder, these four hyperparameters are the threshold values set on the ABDI values and on the NIR band, the buffer pixel distance, and the threshold value of the ABDI pixels included in the buffer, which is denoted hereafter ABDI_bis. The ranges of values taken into account in this analysis are shown in Table 4.4. Note that the threshold value of ABDI_bis depends on that of ABDI. Under no circumstances can the value taken into account for ABDI_bis be greater than the value taken into account for the ABDI threshold. In other words, if the value chosen for ABDI is one, then the possible values for the ABDI_bis threshold are zero or one. Furthermore, the values presented in the rest of the analysis are averaged over the six Sentinel-2 validation images (Table 3.1).

Table 4.4: Summary of value ranges taken into account for model calibration of Sentinel-2 images

| ABDI | NIR | Buffer | ABDI_bis |
|---|---|---|---|
| [0,1,2] | [0,1] | [0,1,2,3,4,5,6,7,8,9,10,20,30,40,50] | [0,1,2] |

The first metric we considered is accuracy. It was calculated for each set of hyperparameters in order to observe the influence of each parameter on this metric (Figure 4.1). Our results show that for the least restrictive parameters (0 - 0 - 0), the accuracy for a zero-pixel buffer size is 0.1 lower than for the other hyperparameter sets. In addition, this curve decreases more rapidly than the others, reaching a plateau of around 0.2 for a buffer size of 30. The other curves are divided into three groups, based on the value of the ABDI_bis parameter. The first group is characterised by solid lines, representing an ABDI_bis value of zero for which the graph pictures a significant decrease in accuracy with increasing buffer size for these hyperparameter sets.We also observed that the more selective the parameters, the lower the accuracy. The second group, represented by dashed lines and corresponding to an ABDI_bis value of one, sees the accuray of each of these parameter sets slightly decrease for the lower buffer sizes before reaching a plateau at around a buffer size of 30. As with the previous group, the more selective the parameters, the higher the accuracy. Finally, the third group represented by the dotted lines, the most selective group for the ABDI_bis parameter, is stable all along the line with a score of 0.9. Increasing the buffer size has no influence on the model's detection capabilities for this group.

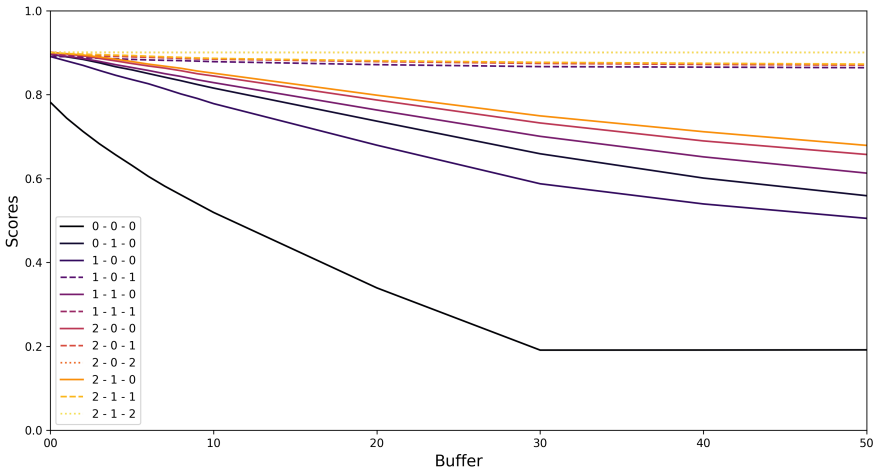



Figure 4.1: Trend in accuracy for each parameter set (ABDI - NIR - ABDI_bis) as a function of buffer value. Solid, dashed and dotted lines respectively represent cases with ABDI_bis = 0, 1 and 2.

Then we considered precision, recall and F1-score and analysed them through 12 graphs (Figure 4.2). We again considered different parameter sets in the form (ABDI - NIR - ABDI_bis) and let the metrics vary according to the buffer size, with values ranging from 0 to 50. Here, we averaged the precision and recall metrics over the six Sentinel-2 validation images. We then calculated the maximum of the F1-score from the average of the two previous metrics.

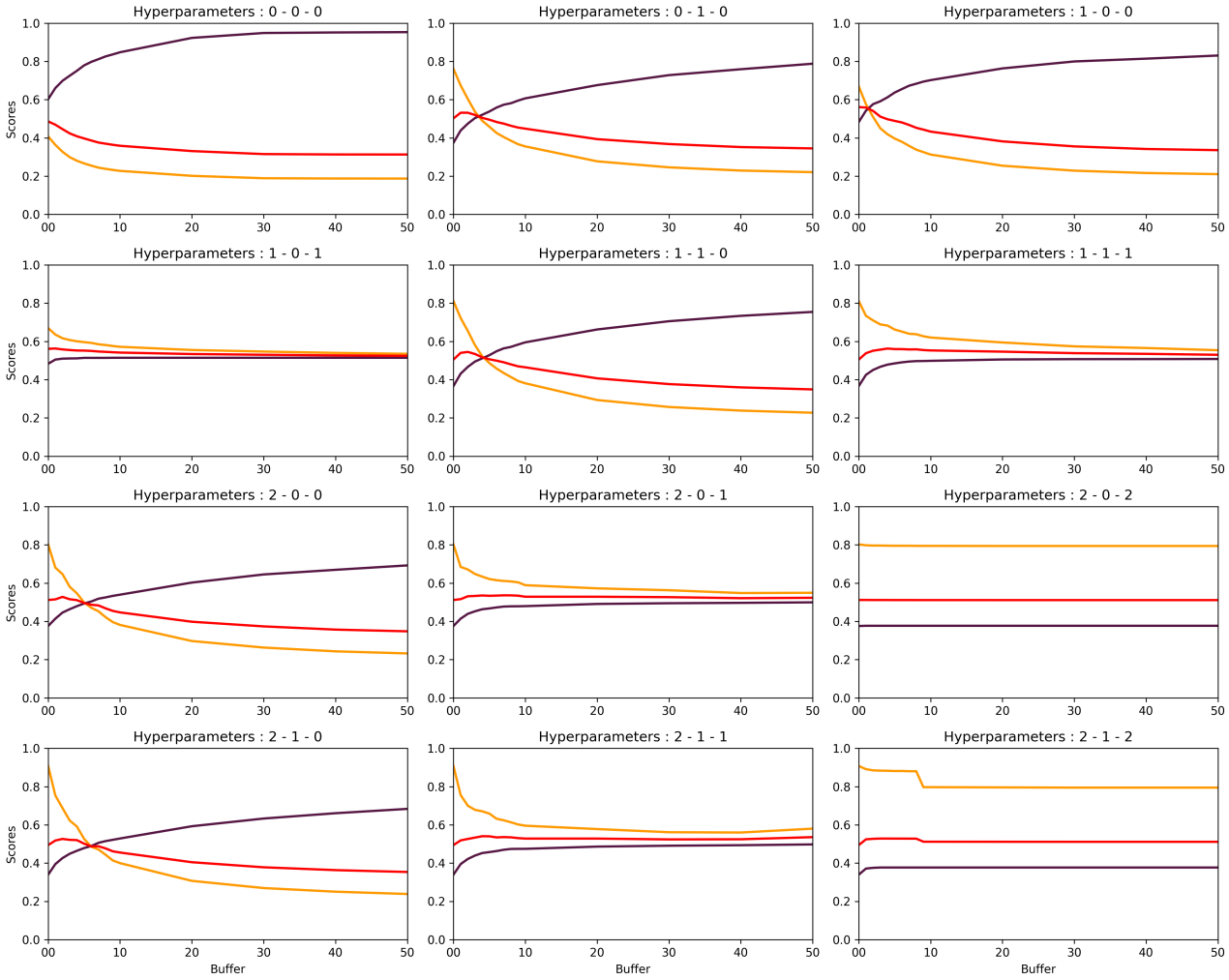



Figure 4.2: Comparison of the evolution of precision (orange), recall (purple) and F1-score (red) as a function of buffer value for each set of parameters (ABDI - NIR - ABDI_bis).

For the first graph (0-0-0), there is no intersection between the precision and recall curves with a decreasing trend for the F1-score. Six of the other graphs (1-0-1, 1-1-1, 2-0-1, 2-0-2, 2-1-1 and 2-1-2) also show no intersection between the precision and recall curves. However, in contrast to the first graph, the trend of the F1-score curve for these other six graphs is slightly increasing before reaching a plateau around 0.5. This stabilization induces F1-score maximisation, leading to a trade-off between precision and recall to maximise both for each parameter set. In the last five graphs (0-1-0, 1-0-0, 1-1-0, 2-0-0 and 2-1-0), there is an intersection between the precision and recall curves. However, as mentionend before, the maximum F1-score value does not lie at the intersection of the two curves, but at buffer values two or three pixels smaller. Finally, all the metrics for each plot generally reach a plateau for buffer sizes between 10 and 20 pixels, while the maximum F1-score values are reached between buffer values of zero and ten, but cannot be precisely determined. To this end, the maximum F1-score values for each hyperparameter set (ABDI - NIR - ABDI_bis) and the associated buffer values are displayed hereafter in Table 4.5.

Table 4.5: Highest F1-score for each hyperparameter set (ABDI - NIR - ABDI_bis) and its associated buffer value.

| **Hyperparameters (ABDI - NIR - ABDI_bis)** | **Buffer** | **F1-score** |
|---|---|---|
| 0 - 0 - 0 | 00 | 0.4862 |
| 0 - 1 - 0 | 01 | 0.5316 |
| 1 - 0 - 0 | 00 | 0.5612 |
| 1 - 0 - 1 | 02 | 0.5627 |
| 1 - 1 - 0 | 02 | 0.5448 |
| 1 - 1 - 1 | 04 | 0.5625 |
| 2 - 0 - 0 | 02 | 0.5279 |
| 2 - 0 - 1 | 08 | 0.536 |
| 2 - 0 - 2 | 01 | 0.5118 |
| 2 - 1 - 0 | 02 | 0.5262 |
| 2 - 1 - 1 | 04 | 0.5404 |
| 2 - 1 - 2 | 03 | 0.5282 |

This table shows maximum F1-scores around 0.5 for each hyperparameter set. The least restrictive set of parameters is the one with the lowest F1-scores while maximum F1-score values ($>$0.56) are achieved for the hyperparameter sets 1-0-0, 1-0-1 and 1-1-1. Moreover, all these maximum values correspond to buffer size values between zero and eight. The impact of buffer size therefore seems to have an impact on the first iterations.

A final tool for choosing the right combination of hyperparameters is the ROC curve (Figure 4.3). It provides a compromise between the true positive rate (TPR) and the false positive rate (FPR). As for Figure 4.2, the results of the ROC curves are presented in 12 graphs where each set of hyperparameters varies according to the fourth hyperparameter, i.e. buffer size. In addition, the dashed line indicates the performance of a random classifier. If one of the points lies below this line, it means that a random classifier performs better than the model used.

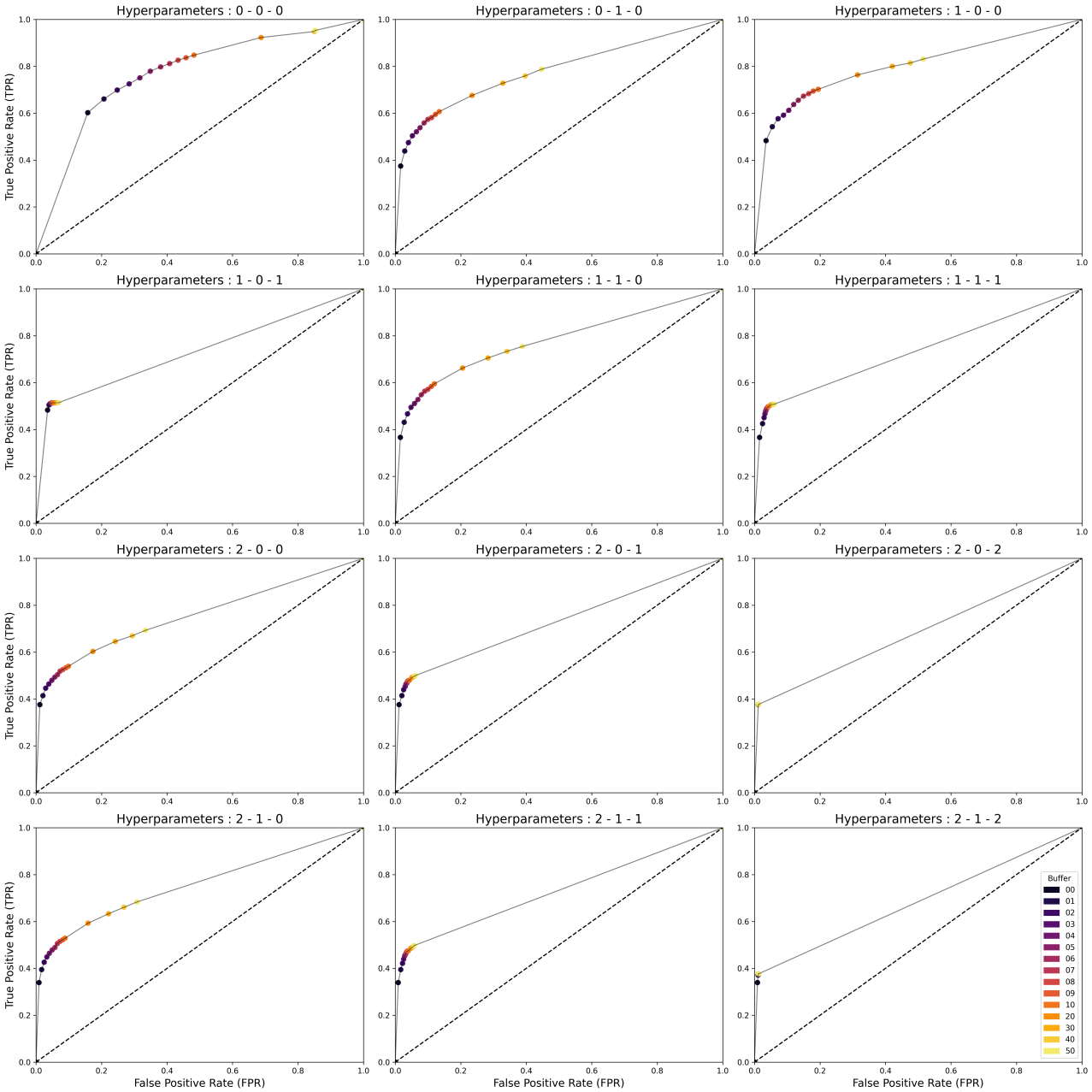



Figure 4.3: ROC curves as a function of different buffer sizes for each parameter set (ABDI - NIR - ABDI_bis).

This figure allow to provide the final information on the choice of hyperparameters. Firstly, none of the 12 graphs show points below the line corresponding to the performance of a random classifier, which means that whatever parameter values are chosen, the result will be better than if each pixel was classified randomly. Furthermore, the 12 graphs can be separated into three categories according to their shape. The first category includes only the first graph (0-0-0),

which has a mild slope. Indeed, for an increase in buffer size, the FPR increases faster than the TPR, which leads to a greater increase in a poor classification than a good one. However, the first value for a zero-pixel buffer size shows a TPR of 0.6 for an FPR of less than 0.2, leading to good prediction of algal blooms pixels without making many errors in their prediction. The second category includes graphs for the following parameter sets: 1-0-1, 1-1-1, 2-0-1, 2-0-2, 2-1-1 and 2-1-2. These graphs see their points grouped together, which means that buffer size has almost no impact on improved detection and never exceeds a TPR of 0.5, meaning that less than one pixel in two is well predicted as an algal bloom. Finally, the last category includes graphs for the following parameters: 0-1-0, 1-0-0, 1-1-0, 2-0-0 and 2-1-0. In these graphs, increasing buffer size rapidly increases TPR relative to FPR before having the opposite effect from a buffer size of ten pixels. The parameter sets in this category show the best results, with each graph showing a TPR value of around 0.6 for an FPR value of less than 0.2. Among this category, the best results are achieved for parameters 1-0-0 and a buffer size between zero and ten, with a peak for a buffer of ten with a TPR of 0.7 and an FPR of 0.2.

## 4.2 Model limitation

It is necessary to find the limits of the model in order to know to what extent it is applicable. With this in mind, a limitation analysis was carried out on Kuwait Bay.We had access to turbidity and surface chlorophyll measurements collected at a buoy located in Kuwait Bay (see location in Figure 4.5). These measurements have a 10-min resolution and cover the entire year 2018. The area is suitable for a limitation analysis due to its high sediment load throughout the year, which is the main obstacle to the detection of algal blooms. The March period was selected because of its high chlorophyll concentration and low turbidity, which may correspond to the appearance of an algal bloom around the buoy, as shown in Figure 4.4.

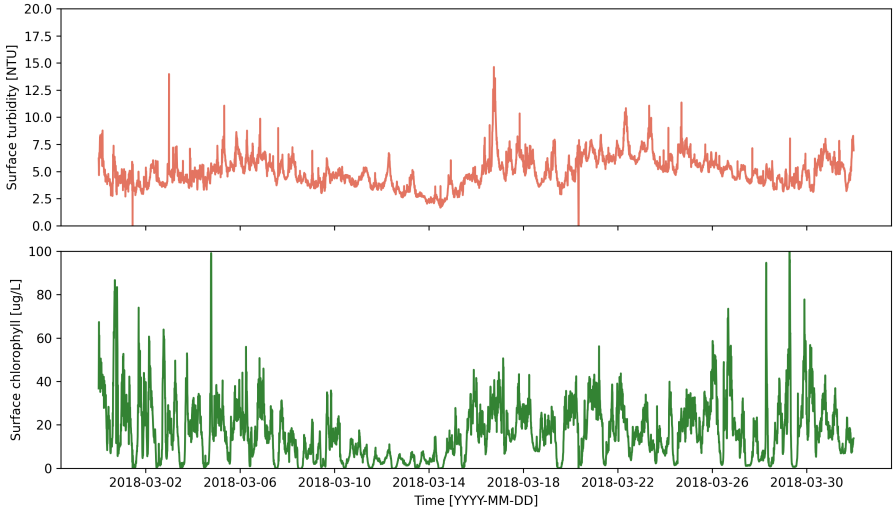



Figure 4.4: Comparison of surface chlorophyll concentration (green) and turbidity (red) for buoy KISR01 (29.436000° 47.903553°) in central Kuwait Bay.

### 4.2.1 Sentinel-3

We considered an analysis based on Sentinel-3 image processing. As the northern zone of the Bay is heavily loaded with sediments, an analysis of its impact on detection was carried out (Figure 4.5). This northern zone is shown in hatched red on both the figure and the bar chart in Figure 4.5. The graph on the right shows the influence of zone masking on the number of pixels detected, and therefore on the detection of algal blooms. When the area to the north is masked, there is a heterogeneous decrease in the number of pixels above an OTCI value of 2.5, with a decrease of almost 800 pixels for March 8, 2018, but only 20 pixels for March 17, 2018. However, this reduction has no impact on the prediction of days with algal blooms. Only March 1, 2018 is impacted by this difference, the rest of the days remain in the same prediction category with or without the mask applied.

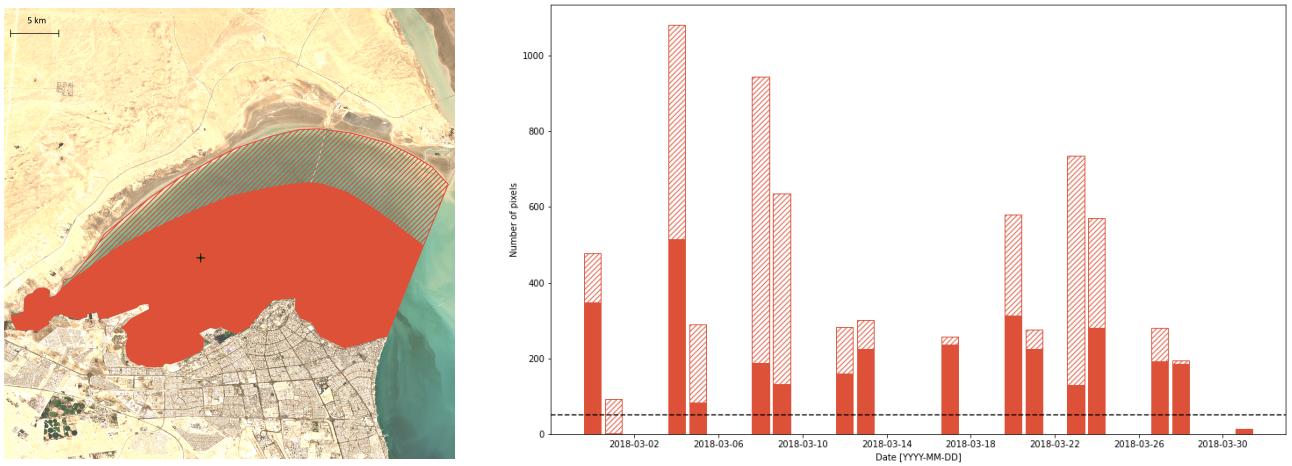



Figure 4.5: Comparison of the number of pixels assigned to the algal bloom category according to two zones (the whole Bay in hatched and the Bay without the northern zone in red).

In addition, a comparison between chlorophyll concentration and OTCI values was carried out (Figure 4.6). To this end, the OTCI value of the pixel containing the buoy (in orange on the graph) was extracted and compared with the reading taken by the buoy (in purple on the graph) for the same day and time as the satellite image. The trend of the two curves on the graph show a first peak at the beginning of the month, followed by a trough and finally a last peak at the end of the month. However, the OTCI trend has less amplitude than the chlorophyll measurements. Moreover, the first peak is more pronounced for OTCI than for chlorophyll. Finally, the second peak is much sharper and appears later for chlorophyll. This peak can be explained by the presence of a high concentration of algae on a small scale near the buoy, making it impossible for OTCI to detect it on a pixel scale. This graph shows a certain similarity between the two metrics, which is to be expected given that both are concerned with chlorophyll.

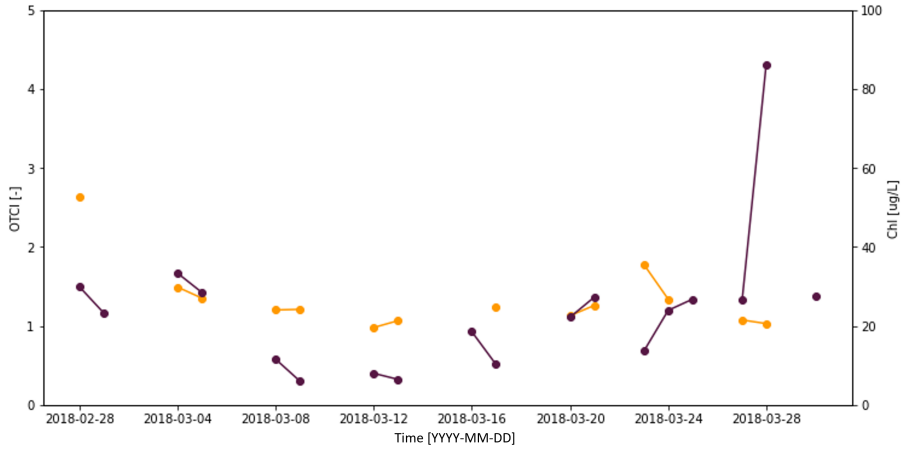



Figure 4.6: Comparison of the chlorophyll concentration at the buoy (purple) and the OTCI value for the pixel containing the buoy (orange).

### 4.2.2 Sentinel-2

We can further refine the analysis of the method's limitations by considering Sentinel-2 images by applying the detection with the hyperparameters derived in Section 4.1 (left panel of Figure 4.7). The pixels predicted by the model are shown in purple and the extent of the algal bloom in hatched red. It is clear to us that the prediction is not good in this case, as the predicted pixels are very small in number and are not even all within the extent of the algal bloom. Some of them are found along the Sheikh Jaber al-Ahmad al-Sabah Bridge and around the boats in the Bay. To better understand this result, it was compared with the ABDI distribution (right panel of Figure 4.7). This figure shows that ABDI fails to capture and delimit the algal bloom, which it is already supposed to do in part before being refined by NIR.




Figure 4.7: On the left, a comparison of the predicted algal bloom zone (purple) and the actual algal bloom zone (hatched). On the right, comparison between the algal bloom zone (red) and the calculated ABDI (the greener the more likely it is to be an algal bloom).

## 4.3 Trend over 2023

With the parameters and the limitations that we have identified, we can now apply our approach over a broader area including the entire Arabian/Persian Gulf, the Strait of Hormuz and the western Gulf of Oman for the entire year 2023.

### 4.3.1 Sentinel-3

We first use the Sentinel-3 results to identify the days during which algal blooms are present (Figure 4.8). We further distinguish Sentinel-3A (orange) and Sentinel-3B (purple), and highlight the most affected months, which are February, April and May in the case of 2023. During certain periods of these months, large algal blooms were observed almost every day, suggesting that a single algal bloom was observed during each of these periods. More specifically, the model predicts that the first algal bloom appeared on January 31 and vanished after July 11, and thus none in the last five months of the year, highlighting a certain seasonality in the appearance of these events in the region. Finally, the use of two satellites enables regular monitoring and more effective temporal detection, as each day is covered by at least one of the two satellites.

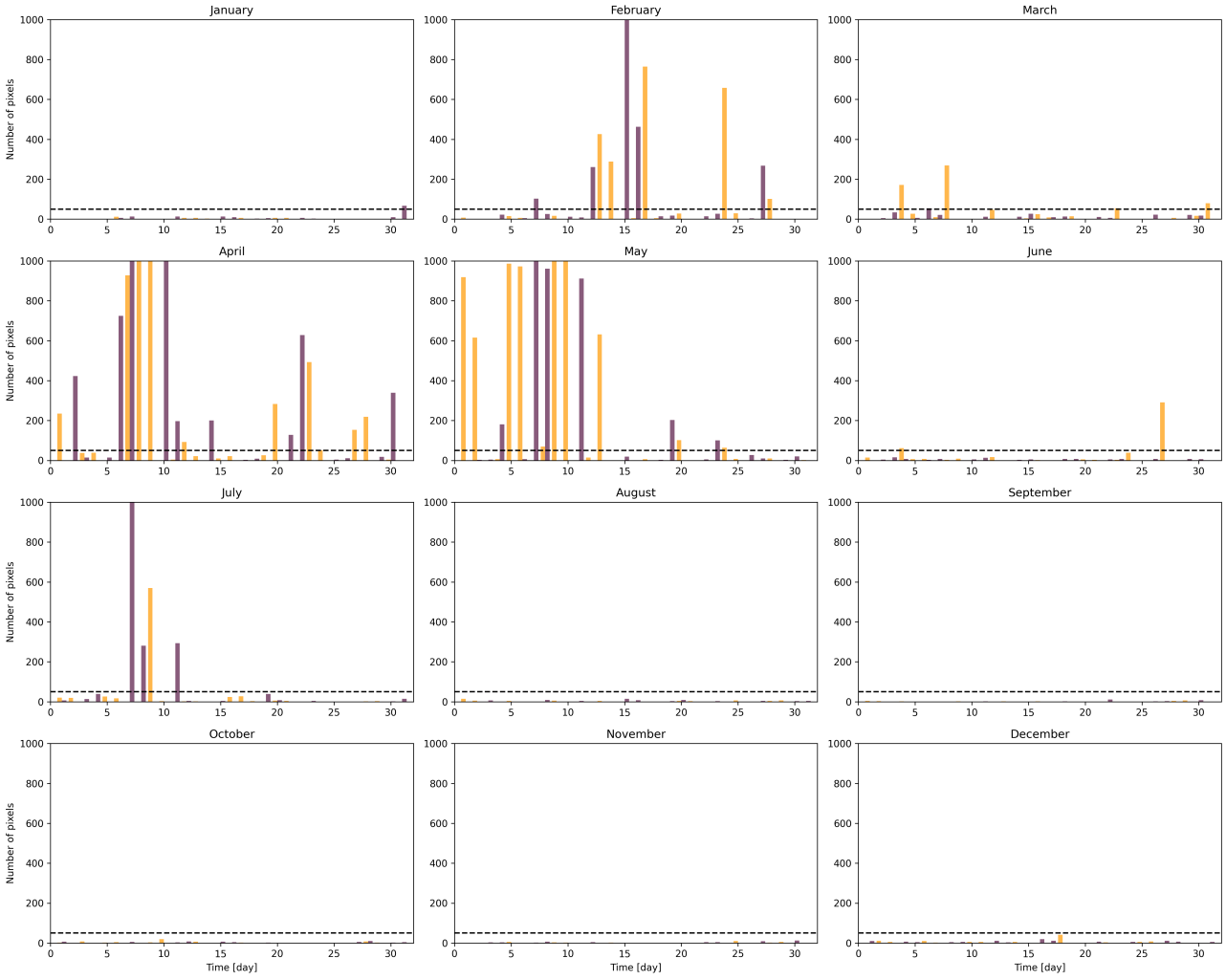



Figure 4.8: Number of pixels whose value is greater than or equal to 2.5 for OTCI month by month for Sentinel-3A (in orange) and Sentinel-3B (in purple). The threshold value of pixels that must be exceeded to predict an algal bloom is shown with a dashed line.

Then, by counting for each month the number of days with an algal bloom, we can obtain a yearly bloom distribution (Figure 4.9). The left panel shows the distribution for the entire area of interest, while the right panel considers only the Gulf. Obviously, there are more blooms in the broader area but these blooms also tend to occur earlier. The bloom peak occurs in April in the broader area and one month later in the Gulf. Moreover, for June and July the number of days is the same for both zones, implying that the algal blooms detected during these months are only localised in the Gulf. In addition, the curve highlights the trend in the number of days with an algal bloom per month over a year, which show a certain seasonality in their outbreaks.

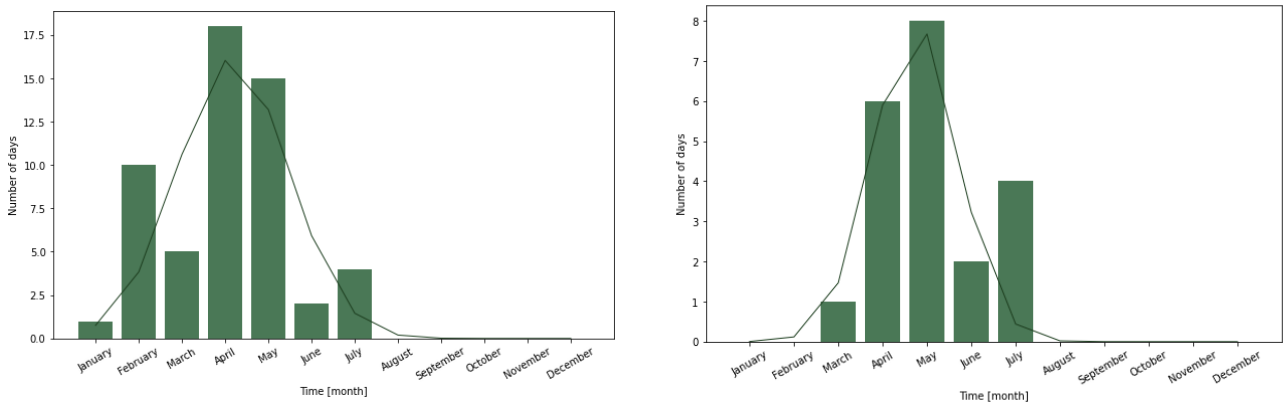



Figure 4.9: Number of days with algal blooms for each month of 2023 for the broad area including the Gulf, the Strait of Hormuz and the western Gulf of Oman (left panel) and only for the Gulf (right panel)

Once all the dates when algal blooms are present have been identified, a map showing the areas most affected during the year 2023 can be produced (Figure 4.10). This map was obtained by adding one to the pixel value if it is considered as an algal bloom for the given date. The map shows that the most affected areas are in the western Gulf of Oman and the Strait of Hormuz. More specifically, it is the outer shores of the meander formed by the connection of the two gulfs where most algal blooms were detected. The center of the Gulf was also affected in 2023, as shown in the close-up view on the left. Of all the pixels in the zone, 0.38% were affected at least once during the year by an algal bloom, representing a surface area of 4,394.7 $km^2$. Of the pixels affected, 88.3% were affected only once, 9.83% twice and only 1.87% more than twice during the whole year. Moreover, this map also shows which desalination plants have been most affected and most exposed to the risk of algal blooms. The three plants in Iran's Hormozgan province and the three in the far northern United Arab Emirates were the most affected by algal blooms.

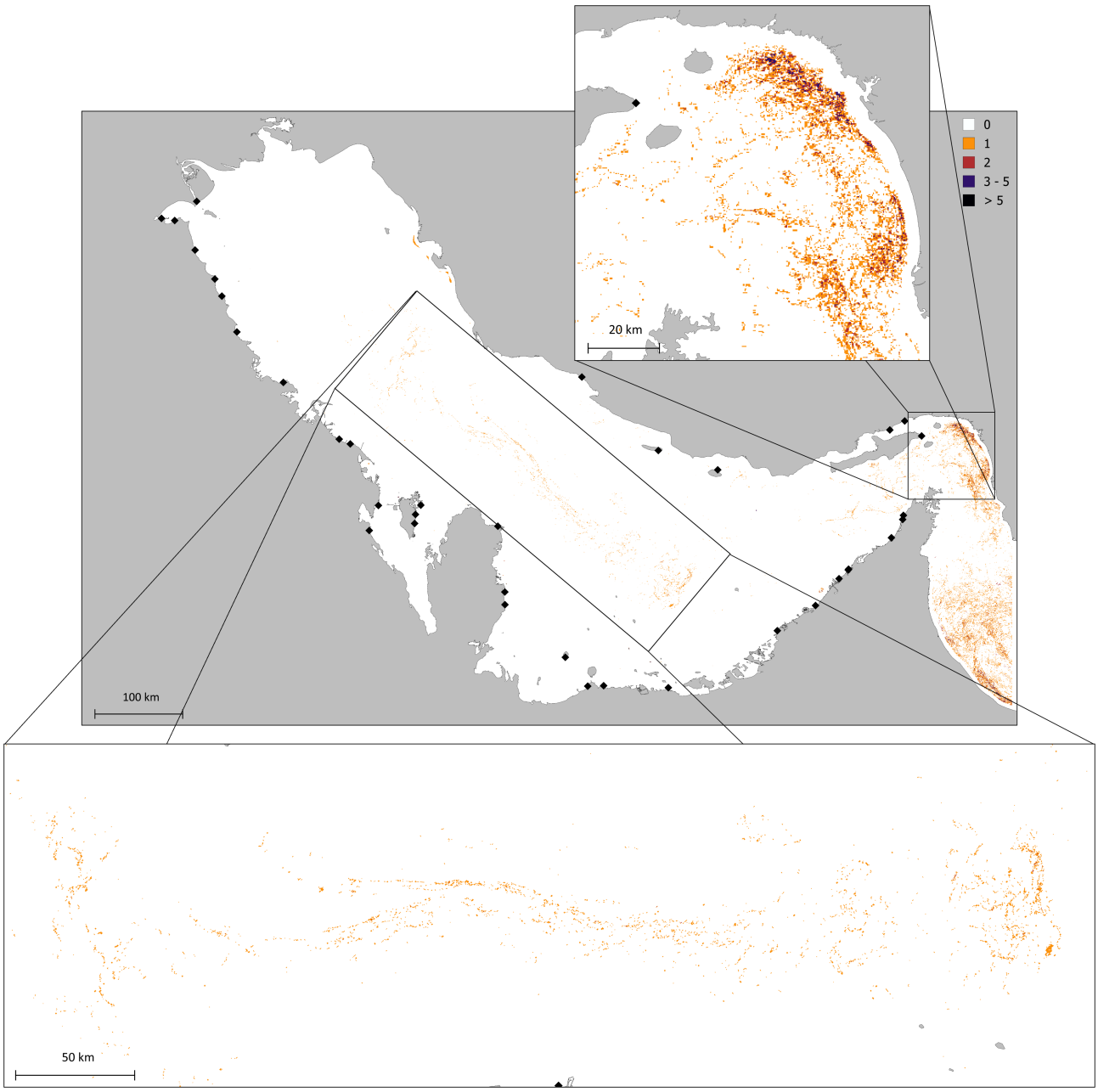



Figure 4.10: Map showing the areas most affected by algal blooms in 2023 and the location of desalination plants in black dots.

### 4.3.2 Sentinel-2

We finally extracted a precise delineation of the areas where algal blooms were detected thanks to Sentinel-2. For the year 2023 and among the dates when an algal bloom was detected, there are not many Sentinel-2 images. However, some dates have usable Sentinel-2 images over the algal bloom area, such as those of February 17, 2023 (Figure 4.11). Thanks to the hyperparameters found during calibration of the model presented above, it is possible to delineate and identify more clearly the extent of the algal bloom (in red on the map). In the area over which Sentinel-2 passed on this date, the model detects an algal bloom with a surface area of 1324.29 $km^2$, mainly located along the coasts of the United Arab Emirates and Oman on the Gulf of Oman. Furthermore, Figure 4.12 allows us to compare what is highlighted on the second close-up view of Figure 4.11 and seems to show that the detection does highlight the main areas of the algal bloom while the clouds on the left-hand side do not interfere with the detection and results.

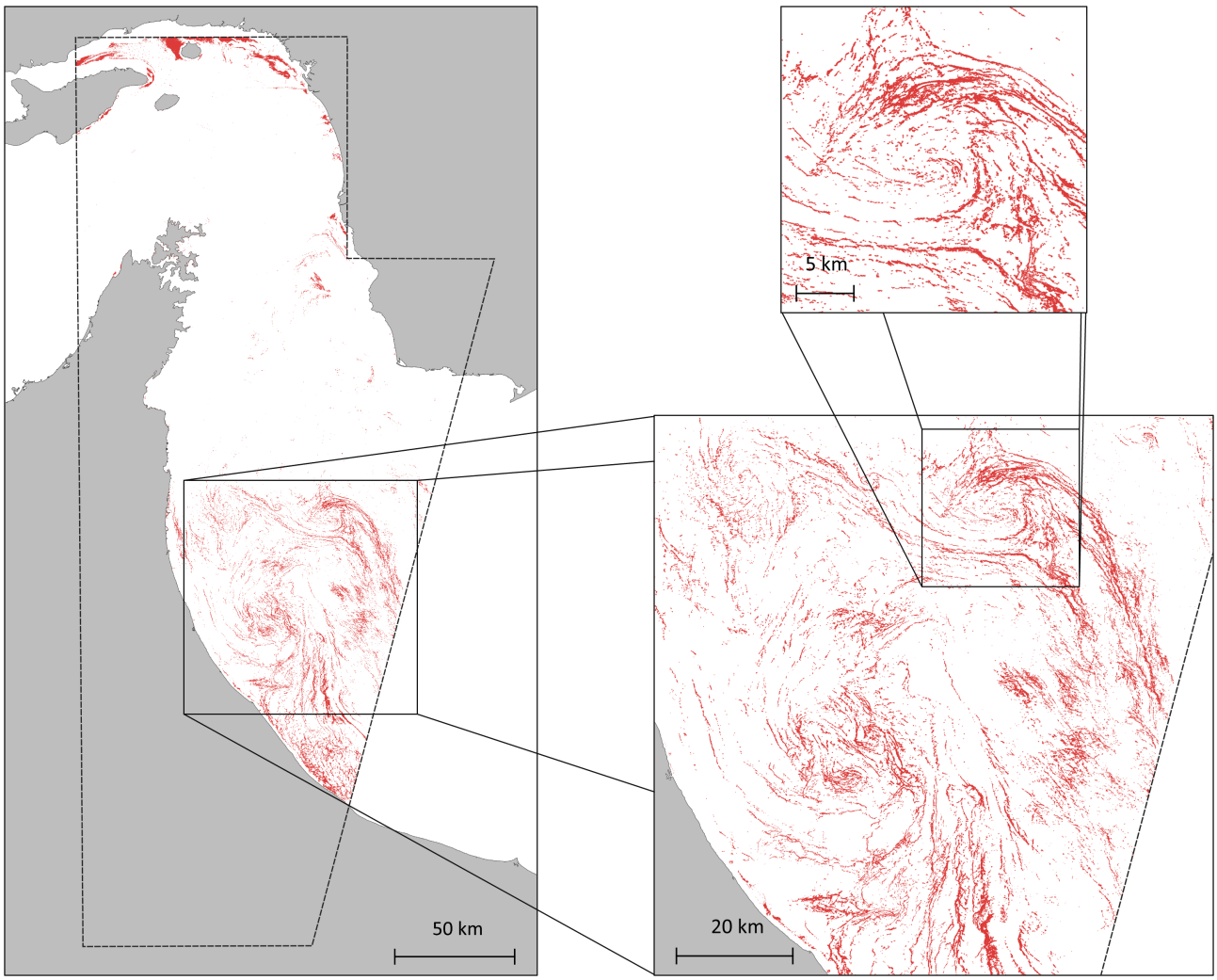



Figure 4.11: Delineation of the February 17, 2023 algal bloom (red) using the Sentinel-2 image method. The last close-up view corresponds to the extent of the image shown in Figure 4.12.

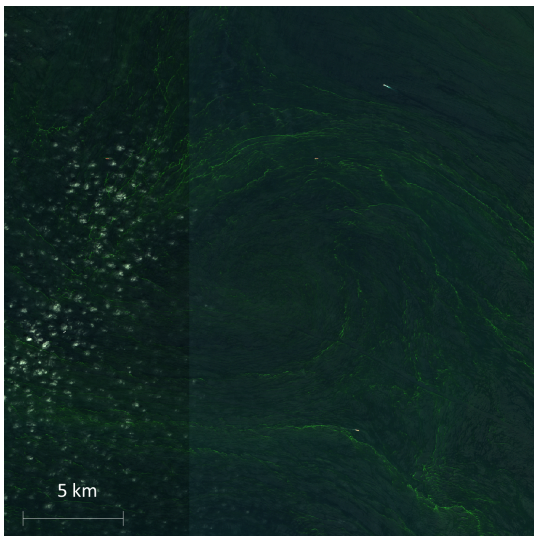



Figure 4.12: Green-spectrum enhanced RGB image of the algal bloom of February 17, 2023. This view corresponds to the last close-up view shown in Figure 4.11.

## 4.4 Link with oil spills

We can finally analyse the impact of oil spills on the formation of algal blooms by assessing whether the presence of an oil spill can trigger or enhance algal blooms. We considered on oil spill detected in the center of the Gulf, east of Qatar, on March 28, 2021 (Figure 4.13). The red area corresponds to the drift of the oil slick as simulated with the oil-spill dispersal model OpenOil ("opendrift model", n.d.) during four days. Finally, the areas shown in green correspond to the algal bloom detected on April 01, 2021 using a Sentinel-3 image and its analysis using the model developed. Unfortunately, no usable Sentinel-2 images were available for these dates to better delimit the algal bloom. Taken together, these results show that most of the algal bloom is found in the area were the oil spill was predicted to be in the four days following its detection. Moreover, part of the algal bloom is to the north, which seems to be not predicted to be affected by the oil spill or its drift.

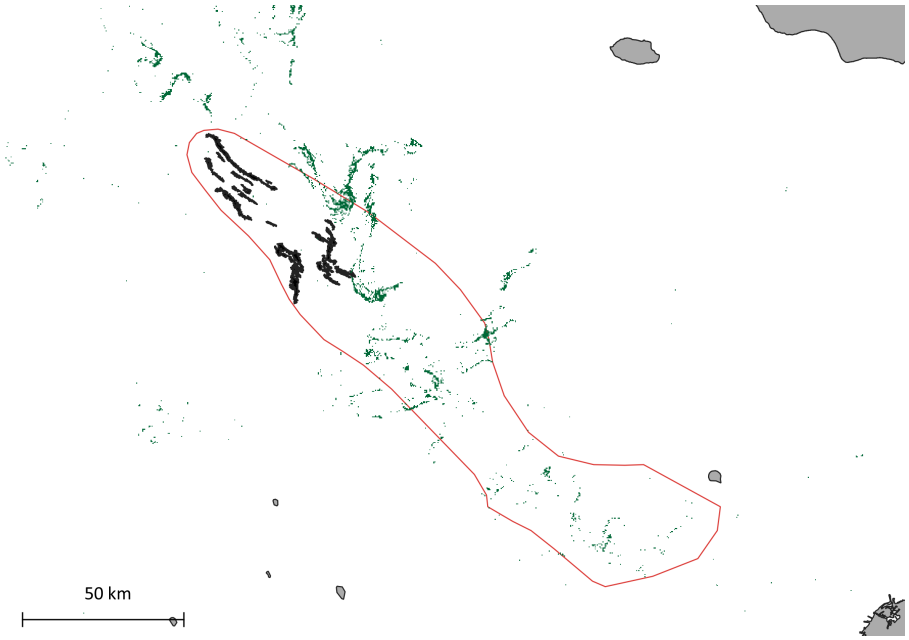



Figure 4.13: Comparison between the oil spill of March 28, 2021 (black) and its extension to April 1, 2021 (red) and the algal bloom observed on April 1, 2021 (green).

CHAPTER

# FIVE

# DISCUSSION

We developed an algal bloom detection algorithm based on a combination of Sentinel-2 and Sentinel-3 images, calibrated the algorithm and applied them to derived a map illustrating where algal bloom were present in 2023 over a large area encompasses the Arabian/Persian Gulf, the Strait of Hormuz and the Gulf of Oman. Moreover, we attempted to assess whether the presence of an oil spill could foster or trigger the development of algal blooms.

The developed and previously described algorithm enables algal bloom daily detection using Sentinel-3 and OTCI and highlights its robustness in detecting all forms of algal blooms, harmless and harmful ones, toxic or non-toxic. The model is therefore not capable of detecting only HABs. However, this extended detection can be considered suitable for the Gulf considering the indirect problems caused by harmless algal blooms as well, such as fish kills and the problems they can cause for desalination plants. This risk assessment is achievable on a daily basis thanks to the use of Sentinel-3 and its relatively short revisit time (one to two images per day for the study area).

The study area is considered at risk when algal blooms are present, as they provoke ecological and/or economic consequences. In fact, the Gulf maritime zone is an essential area for the Gulf surrounding countries, which depend almost entirely on the sea for their economic activities.

The occurrence of algal blooms leads to the death and/or poisoning of fish. This happens either through the release of toxins by HABs or by eutrophication, which deprives the marine environment of light and reduces oxygen levels. These impacts can have substantial economic consequences for the fishing industry, which is the second most important economic activity in the region ("The marine Ecosystem of the Persian Gulf is undercritical environmental condition and destruction", 2024). In the event of a major algal bloom such as the one of 2008-2009, the inability to meet fishing quotas can lead to stock shortages, resulting in economic difficulties. Thanks to our model, the presence or absence of algal blooms can be identified on a daily basis. Additionally, it can identify the most affected area(s), allowing for example fishermen to better plan their fishing trips and maximise their productivity, thus avoiding economic and food crises for local populations.

Secondly, detecting algal blooms would improve the management of desalination plants. Following the detection of a bloom, each plant could draw up an action plan based on its trajctory, thus preventing delayed responses. In addition, the analysis highlights the most impacted plants in 2023, and the map can also be used to identify optimal locations for new plant construction in the future.

Despite the issues caused by algal blooms, it is crucial to emphasize the importance of detection and prediction, rather than eradication. Indeed, algal blooms are essential to many biogeochemical cycles and to the food chain. Controlling them without eliminating them, as ROPME aims, is therefore important. The most effective way to regulate them is to control the human causes mentioned earlier, such as controlling the release of waste and polluted water and oil spills.

The model can also assess the most favorable periods to algal bloom formation. Indeed, a strong seasonality at the beginning of the year was identified, influenced by surface currents in the Indian Ocean and the emergence of algal blooms along the Indian coast (Padmakumar *et al.*, 2012). These algal blooms seem to travel thousands of kilometers to the Gulf, as it was supported by similar observations in other studies (Al Shehhi *et al.*, 2014; Zohdi and Abbaspour, 2019). This seasonality is further influenced by more favorable winter conditions, such as increased phosphate and nitrate concentrations controlled by the high kinetic energy of surface currents, which impact the concentration seasonality (Piontkovski *et al.*, 2019), which rise with the release of industrial and shipping waste. Increased controls and restrictions during these periods could therefore help manage these recurrent events. However, given the region's reliance on oil transport for its economy, limiting this activity in the near future seems challenging. Thus, better management of polluted water releases and deliberate oil spills in the Gulf, alongside increased supervision of oil tanker routes to prevent massive oil releases resulting from incidents for example, are recommended. In fact, it was proven that there is a link between oil spills and the development of algal blooms, oil release into the waters of the Gulf leads to an increase in algal blooms. Preventing these releases would therefore reduce bloom occurrences, without necessarily impeding them.

The used algorithm, and more precisely the model's combination of the ABDI and NIR, also yields promising results. The chosen parameters, 1-0-0-06 (ABDI - NIR - ABDI_bis – buffer value), balance the results of different metrics. They provide good algal bloom prediction while minimising the prediction error. However, our model's F1-score is 0.48, with an accuracy of 0.8, compared to accuracies over 0.95 in other studies (Cao *et al.*, 2021). This discrepancy is due to class imbalance, lowering the precision metric. Nevertheless, the tests carried out show the model's ability to accurately detect algal blooms identified by Sentinel-3. We therefore consider our prediction results to be acceptable and reliable.

Furthermore, the analysis highlights the filamentous structure of algal blooms, showing the model's ease in identifying first converging fronts before extending detection using the buffer values. This result was expected as the model was built around this feature. This also reveals the surface area of blooms, aligning with other studies that report areas ranging from hundreds to thousands of square kilometers, and filaments up to 100 km long (Schleyer and Vardi, 2020; Smith *et al.*, 2014). Accurate bloom delimitation would improve desalination plant management, enhancing water filtering efficiency and drinking water stock management, benefiting all sectors in the affected countries surrounding the Gulf.

Combining data from Sentinel-2 and Sentinel-3 have therefore provided a robust model for spatial and temporal observations, respectively. However, although the results are encouraging, several limitations hinder the algorithm's current usability.

The first limitation is the model's difficulty in detecting algal blooms in turbid zones, as shown in a study on Kuwait Bay. Sentinel-3 detection in these areas requires prior knowledge of bloom presence due to the tendency of colored organic matter inducing overestimation of chlorophyll concentration. This issue makes automatic detection using OTCI impossible without prior bloom knowledge. Additionally, it seems that the mixture in the Kuwait Bay water column visually masks algal blooms (Al-Ghadban and El-Sammak, 2005), preventing Sentinel-2 from accurately identifying the algal bloom convergence fronts. This limitation thus renders the algorithm ineffective for desalination plants in turbid environments such as the three plants located in the Kuwait Bay, and hampers automatic bloom detection in the entire Gulf.

Secondly, the model only detects large algal blooms with relatively high chlorophyll concentrations. Calibration lead to set parameters with OTCI values greater than 2.5 and a pixel number greater than 50 pixels as they are unable to detect relatively small algal blooms and/or those with low concentrations. Despite our determination to choose parameters that allow overestimation, we still tend to underestimate the number of algal blooms. This underestimation could affect emergency plans by relevant authorities in various economic sectors, such as fishing. Indeed, if the extent of algal proliferation is underestimated, it could hinder the effective use of predictive tools that could enable better management of the marine ecosystem.

In addition, satellite image detection is only possible at the water surface. The presence of algal blooms can therefore be identified at the water surface, but not at depth. Sentinel-2's strength lies in detecting convergence fronts, which do not represent the entire bloom. This poses a problem for desalination plants as the verticality of the algal bloom is not considered. Indeed, if desalination plants were to pump water from deep in the sea with the aim of minimising the quantity of algae, they could find themselves confronted with surprising quantities of algae, which could once again disrupt the overlying socio-economic chain. Two potential solutions are obtaining field data for combined observations, and using 3D simulation models of marine dynamics. These 3D models could assess concentrations at depth using a dispersion of points based on the filament detection, which would therefore improve spatial bloom detection and model robustness.

Furthermore, the model is currently not specific to HABs. To address this, it is possible to integrate a species prediction model as demonstrated in other studies (Ghatkar *et al.*, 2019). This model uses algae absorbance to determine species and bloom category (harmless or harmful), providing valuable decision-making information. For fishing for example, bloom category impacts operations: harmless blooms may reduce yield but remain fishable, whereas toxin releasing HABs pose human health risks, making the area unsuitable. The same impacts can be expected for desalination plant operations.

Finally, there are several intrinsic limitations to optical remote sensing and preprocessing. Sentinel-2's revisit time hinders accurate algal bloom identification. In 2023 for example, of the 55 days retrieved by Sentinel-3, only three coincided with Sentinel-2 observations in algal bloom zones. Adding other high-resolution optical sensor satellites to the model would enhance the accuracy of algal bloom detection. Including Landsat 7 (15-60 m resolution with a 16-day revisit time), Landsat 8 (15-100 m resolution with a 16-day revisit time) or Landsat 9 (15 m resolution with a 16-day revisit time) would significantly improve temporal and spatial coverage, thereby increasing the availability for precise detection.

Another limitation is the time-consuming SNAP preprocessing required when using Sentinel-3 images. Additionally, these preprocessing steps can vary with upstream product updates, preventing full automation. In contrast, Sentinel-2 requires no such preprocessing, making it easier to handle. Aligning Sentinel-3's product structure with Sentinel-2's could simplify and expedite model development, allowing partial automation.

The final limitation is the impact of cloud presence on observations. Although this study area experiences less cloud cover annually than other regions, this constraint remains. Clouds can lead to detection errors or obscure algal blooms, making them undetectable by the model if they are not adequately masked. It is also possible that if the mask is well applied, any algal blooms under these clouds will not be detected by our model. In the future, combining multiple satellites to obtain images at different times of day, as previously described with the landsat satellites, could mitigate this issue and improve event monitoring.

CHAPTER

# SIX

# CONCLUSION

Algal blooms, both harmless and harmful, threaten marine environments, economies, and public health, particularly in the frequently affected Arabian/Persian Gulf. With the region's reliance on seawater desalination due to limited freshwater sources, establishing an early warning system for algal blooms is essential for the efficient management of desalination plants and related economic activities. This work aimed to establish an algal bloom detection system for the Arabian/Persian Gulf and explore possibilities for its automation.

To this end, we innovatively incorporated the OTCI for algal bloom detection to leverage its relatively low revisit time for more regular monitoring. This index was coupled with a model based on Sentinel-2 images, linking the ABDI to the NIR for more accurate detection of algal blooms.

On one hand, Sentinel-3 demonstrated its ability for daily algal bloom detection using OTCI, identifying 55 days with algal blooms in 2023 across the Gulf, including the Strait of Hormuz and the western Gulf of Oman. This highlighted the seasonality of algal blooms, influenced by water surface currents and other factors in the area. In addition, it allowed us to create a risk map for 2023. This map identified the most affected desalination plants and suitable areas for future plant installations. Furthermore, the model suggested a possible link between oil spills and algal blooms, which could expand the possibilities and interests in detecting and predicting these events.

On the other hand, the combination of ABDI and NIR enabled more precise surface algal bloom trend delineation. In future studies, it will be interesting to determine the 3D structure of algal blooms for better spatial prediction, which would help desalination plants determine safe water extraction depths for human consumption. However, the full algorithm showed limitations in turbid zones, requiring further refinement and calibration to improve detection and make it applicable to the entire Gulf.

These various analyses laid the foundation for an automatic detection model for the Gulf. While full automation is currently not feasible due to inherent model limitation, Sentinel-3 images already offer strong daily monitoring capabilities despite processing challenges (such as the images' structure and stage of development). The use of Sentinel-2 alone is insufficient for precision acquisition, which is why we recommend incorporating additional satellites in future work. Finally, despite its limitations, this model serves as an interesting proof-of-concept for further studies in this field.

# BIBLIOGRAPHY


Ahn, Y.-H., Shanmugam, P., Ryu, J.-H., & Jeong, J.-C. (2006). Satellite detection of harmful algal bloom occurrences in korean waters. *Harmful Algae*, *5*(2), 213–231. https://doi.org/10.1016/j.hal.2005.07.007

Al Senafi, F., & Anis, A. (2015). Shamals and climate variability in the northern arabian/persian gulf from 1973 to 2012. *International Journal of Climatology*, *35*(15), 4509–4528. https://doi.org/10.1002/joc.4302

Al Shehhi, M. R., Gherboudj, I., & Ghedira, H. (2014). An overview of historical harmful algae blooms outbreaks in the arabian seas. *Marine Pollution Bulletin*, *86*(1), 314–324. https://doi.org/10.1016/j.marpolbul.2014.06.048

Al-Ghadban, A. N., & El-Sammak, A. (2005). Sources, distribution and composition of the suspended sediments, kuwait bay, northern arabian gulf. *Journal of Arid Environments*, *60*(4), 647–661. https://doi.org/10.1016/j.jaridenv.2004.07.017

Alosairi, Y., van Maren, D. S., Scheel, F., Smits, B. P., & Al-Khaldi, M. S. (2024). Aeolian desert dust as a primary estuarine sediment source: Fine sediment transport and dynamics in the northern arabian/persian gulf. *Estuarine, Coastal and Shelf Science*, *302*, 108771. https://doi.org/10.1016/j.ecss.2024.108771

Alvarez-Vanhard, E., Corpetti, T., & Houet, T. (2021). UAV & satellite synergies for optical remote sensing applications: A literature review. *Science of Remote Sensing*, *3*, 100019. https://doi.org/10.1016/j.srs.2021.100019

Anderson, D. M. (2001, January 1). Phytoplankton blooms. In J. H. Steele (Ed.), *Encyclopedia of ocean sciences* (pp. 2179–2192). Academic Press. https://doi.org/10.1006/rwos.2001.0050

Anderson, D. M., Boerlage, S. F. E., & Dixon, M. B. (2017). *Harmful algal blooms (HABs) and desalination: A guide to impacts, monitoring and management.* (Report). Intergovernmental Oceanographic Commission of UNESCO, https://doi.org/10.25607/OBP-203

Ansell, A., Barnes, M., Gibson, R. N., & Gibson, R. N. (1998, October 9). *Oceanography and marine biology: An annual review. volume 36*. CRC Press.

Bartlett, K. (2004). Dust storm forecasting for al udeid AB, qatar: An empirical analysis, 103.

Cao, M., Qing, S., Jin, E., Hao, Y., & Zhao, W. (2021). A spectral index for the detection of algal blooms using sentinel-2 multispectral instrument (MSI) imagery: A case study of hulun lake, china. *International Journal of Remote Sensing*, *42*(12), 4514–4535. https://doi.org/10.1080/01431161.2021.1897186

Chylek, P., Robinson, S., Dubey, M. K., King, M. D., Fu, Q., & Clodius, W. B. (2006). Comparison of near-infrared and thermal infrared cloud phase detections. *Journal of Geophysical Research: Atmospheres*, *111*. https://doi.org/10.1029/2006JD007140

Coles, S. L. (2003). Coral species diversity and environmental factors in the arabian gulf and the gulf of oman: A comparison to the indo-pacific region. http://repository.si.edu/xmlui/handle/10088/5825

*Copernicus browser* [Copernicus browser]. (n.d.). https://browser.dataspace.copernicus.eu/

Cullen, J., Ciotti, A., Davis, R., & Lewis, M. (1997). Optical detection and assessment of algal blooms. *Limnology and Oceanography*, *42*, 1223–1239. https://doi.org/10.4319/lo.1997.42.5_part_2.1223

Dolatabadi, F., Attaran-Fariman, G., & Loghmani, M. (2021). Bloom occurrence and phylogeny of gonyaulax polygramma (dinophyceae) isolated from south east coast of iran (oman sea). *Iranian Journal of Fisheries Sciences*, *20*(6). https://doi.org/10.22092/ijfs.2021.125501

Duncan, G. (2018, May 16). *Saadiyat beach safe for swimmers again after red tide closure* [The national]. https://www.thenationalnews.com/uae/saadiyat-beach-safe-for-swimmers-again-after-red-tide-closure-1.731010

Earth Science Data Systems, N. (2024, May 16). *SeaWiFS | earthdata*. https://www.earthdata.nasa.gov/sensors/seawifs

Environment, U. N. (2017, August 17). *ROPME sea area* [UNEP - UN environment programme]. http://www.unep.org/ropme-sea-area

Fan, X., Liu, Y., Wu, G., & Zhao, X. (2020). Compositing the minimum NDVI for daily water surface mapping.

Ghatkar, J. G., Singh, R. K., & Shanmugam, P. (2019). Classification of algal bloom species from remote sensing data using an extreme gradient boosted decision tree model. *International Journal of Remote Sensing*, *40*(24), 9412–9438. https://doi.org/10.1080/01431161.2019.1633696

GISGeography. (2015, October 6). *World geodetic system (WGS84)* [GIS geography]. https://gisgeography.com/wgs84-world-geodetic-system/

GISGeography. (2017, December 10). *What is bilinear interpolation?* [GIS geography]. https://gisgeography.com/bilinear-interpolation-resampling/

Gobler, C. J. (2020). Climate change and harmful algal blooms: Insights and perspective. *Harmful Algae*, *91*, 101731. https://doi.org/10.1016/j.hal.2019.101731

Hallegraeff, G. M. (2021). *Global harmful algal bloom: Status report 2021 : A scientific summary for policy makers*. UNESCO. Paris, France. http://hdl.handle.net/1834/42355

Heller, U., Guerrin, F., Struss, P., & Roque, W. (n.d.). A qualitative modeling approach to algal bloom prediction.

*Homepage | copernicus*. (n.d.). https://www.copernicus.eu/en

Hoo, Z. H., Candlish, J., & Teare, D. (2017). What is an ROC curve? *Emergency medicine journal: EMJ*, *34*(6), 357–359. https://doi.org/10.1136/emermed-2017-206735

Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, *113*(10), 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012

Hu, C., Muller-Karger, F. E., Taylor, C. (, Carder, K. L., Kelble, C., Johns, E., & Heil, C. A. (2005). Red tide detection and tracing using MODIS fluorescence data: A regional example in SW florida coastal waters. *Remote Sensing of Environment*, *97*(3), 311–321. https://doi.org/10.1016/j.rse.2005.05.013

*Image of the day | copernicus*. (n.d.). https://www.copernicus.eu/en/media/image-day?combine=bloom

John, V., Coles, S., & Abozed, A. (1990). Seasonal cycles of temperature, salinity and water masses of the western arabian gulf. *Oceanologica Acta*. https://www.semanticscholar.org/paper/Seasonal-cycles-of-temperature%2C-salinity-and-water-John-Coles/c1a81a202e90f768aec7f1a1fda94b4293771816

Khan, N. Y., Munawar, M., & Price, A. R. G. (2002, July). *The gulf ecosystem health and sustainability*. Aquatic Ecosystem Health, & Management Society.

Kumar, T. (n.d.). A theory based on conversion of RGB image to gray. *International Journal of Computer Applications*, *7*.

*Lake pyramid algal bloom | copernicus*. (n.d.). https://www.copernicus.eu/en/media/image-day-gallery/lake-pyramid-algal-bloom

*Landsat 8 | landsat science*. (2021, November 30). https://landsat.gsfc.nasa.gov/satellites/landsat-8/

Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, *91*, 216–231. https://doi.org/10.1016/j.patcog.2019.02.023

The marine ecosystem of the persian gulf is undercritical environmental condition and destruction. (2024). *4th INTERNATIONAL CONFERENCE ON BIOLOGICAL RESEARCH AND APPLIED SCIENCE*. https://doi.org/10.37962/IBRAS/2024/84-86

Mcneil, S. (n.d.). *Growing algae bloom in arabian sea tied to climate change*. https://phys.org/news/2017-03-algae-bloom-arabian-sea-tied.html

Michael Reynolds, R. (1993). Physical oceanography of the gulf, strait of hormuz, and the gulf of oman—results from the *Mt Mitchell* expedition. *Marine Pollution Bulletin*, *27*, 35–59. https://doi.org/10.1016/0025-326X(93)90007-7

*Oil production* [Our world in data]. (n.d.). https://ourworldindata.org/grapher/oil-production-by-country

*OLCI applications*. (n.d.). https://sentiwiki.copernicus.eu/web/olci-applications

*Opendrift model*. (n.d.). https://opendrift.github.io/autoapi/opendrift/models/openoil/index.html

O'Reilly, J. E., & Werdell, P. J. (2019). Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6. *Remote Sensing of Environment*, *229*, 32–47. https://doi.org/10.1016/j.rse.2019.04.021

Padmakumar, K. B., Menon, N. R., & Sanjeevan, V. N. (2012). Is occurrence of harmful algal blooms in the exclusive economic zone of india on the rise? *International Journal of Oceanography*, *2012*(1), 263946. https://doi.org/10.1155/2012/263946

Piontkovski, S. A., Hamza, W. M., Al-Abri, N. M., Al-Busaidi, S. S. Z., & Al-Hashmi, K. A. (2019). A comparison of seasonal variability of arabian gulf and the sea of oman pelagic

ecosystems. *Aquatic Ecosystem Health & Management*, *22*(2), 108–130. https://doi.org/10.1080/14634988.2019.1621133

*Preprocessing sentinel-2 images — eo-toolbox 0.1 documentation*. (n.d.). https://eo-toolbox-rtd.readthedocs.io/en/latest/sentinel_2/s2_prepro.html#3.-Atmospheric-correction-&-Cloud-screening

Radočaj, D., Obhođaš, J., Jurišić, M., & Gašparović, M. (2020). Global open data remote sensing satellite missions for land monitoring and conservation: A review. *Land*, *9*(11), 402. https://doi.org/10.3390/land9110402

Rao, P. G., Hatwar, H. R., Al-Sulaiti, M. H., & Al-Mulla, A. H. (2003). Summer shamals over the arabian gulf. *Weather*, *58*(12), 471–478. https://doi.org/10.1002/wea.6080581207

*Red tide: Karenia brevis* [Thompson earth systems institute]. (2018, July 15). https://www.floridamuseum.ufl.edu/earth-systems/blog/red-tide-karenia-brevis/

Richardson, L. L. (1996). Remote sensing of algal bloom dynamics. *BioScience*, *46*(7), 492–501. https://doi.org/10.2307/1312927

Richlen, M. L., Morton, S. L., Jamali, E. A., Rajan, A., & Anderson, D. M. (2010). The catastrophic 2008–2009 red tide in the arabian gulf region, with observations on the identification and phylogeny of the fish-killing dinoflagellate *Cochlodinium polykrikoides*. *Harmful Algae*, *9*(2), 163–172. https://doi.org/10.1016/j.hal.2009.08.013

*S2 mission*. (n.d.). https://sentiwiki.copernicus.eu/web/s2-mission

*S3 mission*. (n.d.). https://sentiwiki.copernicus.eu/web/s3-mission

Sanderman, J., Savage, K., & Dangal, S. R. (2020). Mid-infrared spectroscopy for prediction of soil health indicators in the united states. *Soil Science Society of America Journal*, *84*(1), 251–261. https://doi.org/10.1002/saj2.20009

Schleyer, G., & Vardi, A. (2020). Algal blooms. *Current Biology*, *30*(19), R1116–R1118. https://doi.org/10.1016/j.cub.2020.07.011

Schmid, D. (2020). Le Golfe, de la géopolitique à la politique tout court. *Politique étrangère*, *Printemps*(1), 9–17. https://doi.org/10.3917/pe.201.0009

Sellner, K. G., Doucette, G. J., & Kirkpatrick, G. J. (2003). Harmful algal blooms: Causes, impacts and detection. *Journal of Industrial Microbiology and Biotechnology*, *30*(7), 383–406. https://doi.org/10.1007/s10295-003-0074-9

*Sentinel-2 - data products - sentinel handbook* [Sentinel online]. (n.d.). https://copernicus.eu/missions/sentinel-2/data-products

*Sentinel-2 - satellite description - sentinel online* [Sentinel online]. (n.d.). https://copernicus.eu/missions/sentinel-2/satellite-description

*Sentinel-3 - data products - sentinel online - sentinel online*. (n.d.). https://sentinel.esa.int/en/web/sentinel/missions/sentinel-3/data-products

*Sentinel-3 - mission summary - sentinel online* [Sentinel online]. (n.d.). https://copernicus.eu/missions/sentinel-3/overview/mission-summary

Shankar, D., Vinayachandran, P., & Unnikrishnan, A. (2002). The monsoon currents in the north indian ocean. *Progress in Oceanography*, *52*(1), 63–120. https://doi.org/10.1016/S0079-6611(02)00024-1

Smith, W. O., Liu, X., Tang, K. W., DeLizo, L. M., Doan, N. H., Nguyen, N. L., & Wang, X. (2014). Giantism and its role in the harmful algal bloom species *Phaeocystis globosa*. *Deep Sea Research Part II: Topical Studies in Oceanography*, *101*, 95–106. https://doi.org/10.1016/j.dsr2.2012.12.005

Steidle Neto, A. J., Lopes, D. C., Pinto, F. A. C., & Zolnier, S. (2017). Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. *Biosystems Engineering*, *155*, 124–133. https://doi.org/10.1016/j.biosystemseng.2016.12.008

*STEP – science toolbox exploitation platform*. (2024, May 16). https://step.esa.int/main/

Thoppil, P. G., & Hogan, P. J. (2010). Persian gulf response to a wintertime shamal wind event. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*(8), 946–955. https://doi.org/10.1016/j.dsr.2010.03.002

*An upended ecosystem in the arabian sea*. (2021, August 5). https://earthobservatory.nasa.gov/images/148655/an-upended-ecosystem-in-the-arabian-sea

Vaughan, G. O., Al-Mansoori, N., & Burt, J. A. (2019, January 1). Chapter 1 - the arabian gulf. In C. Sheppard (Ed.), *World seas: An environmental evaluation (second edition)* (pp. 1–23). Academic Press. https://doi.org/10.1016/B978-0-08-100853-9.00001-4

Wang, M., Jiang, L., Mikelsons, K., & Liu, X. (2021). Satellite-derived global chlorophyll-*a* anomaly products. *International Journal of Applied Earth Observation and Geoinformation*, *97*, 102288. https://doi.org/10.1016/j.jag.2020.102288

*What is a harmful algal bloom? | national oceanic and atmospheric administration*. (n.d.). https://www.noaa.gov/what-is-harmful-algal-bloom

Xu, D., Pu, Y., Zhu, M., Luan, Z., & Shi, K. (2021). Automatic detection of algal blooms using sentinel-2 MSI and landsat OLI images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 8497–8511. https://doi.org/10.1109/JSTARS.2021.3105746

Yu, S., Gong, F., He, X., Bai, Y., Zhu, Q., Wang, D., & Chen, P. (n.d.). Satellite views of the massive algal bloom in the persian gulf and the gulf of oman during 2008-2009.

Zhao, J., & Ghedira, H. (2014). Monitoring red tide with satellite imagery and numerical models: A case study in the arabian gulf. *Marine Pollution Bulletin*, *79*(1), 305–313. https://doi.org/10.1016/j.marpolbul.2013.10.057

Zhao, J., Temimi, M., & Ghedira, H. (2015). Characterization of harmful algal blooms (HABs) in the arabian gulf and the sea of oman using MERIS fluorescence data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *101*, 125–136. https://doi.org/10.1016/j.isprsjprs.2014.12.010

Zohdi, E., & Abbaspour, M. (2019). Harmful algal blooms (red tide): A review of causes, impacts and approaches to monitoring and prediction. *International Journal of Environmental Science and Technology*, *16*(3), 1789–1806. https://doi.org/10.1007/s13762-018-2108-x

Zwietering, M. H., Jongenburger, I., Rombouts, F. M., & Van 'T Riet, K. (1990). Modeling of the bacterial growth curve. *Applied and Environmental Microbiology*, *56*(6), 1875–1881. https://doi.org/10.1128/aem.56.6.1875-1881.1990

APPENDIX

# A

# TABLE OF IMAGES WITHOUT ALGAL BLOOMS USED FOR VALIDATION

Table A.1: Summary of satellite images used for the validation process without algal blooms.

| Satellite | Location | Date |
|---|---|---|
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2017-02-24 |
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2019-06-19 |
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2020-10-17 |
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2020-12-10 |
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2022-08-05 |
| S3A_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2023-06-25 |
| S3B_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2021-06-27 |
| S3B_L1_EFR | Gulf & Strait of Hormuz & Gulf of Oman | 2022-04-13 |

# Integrated system for detecting algal blooms in the Arabian/Persian Gulf: A multi-sensor approach using the Sentinel-2 and Sentinel-3 satellites data

Julien Thonnard

Harmless and harmful algal blooms present significant threats to marine environments, economies, and the health of exposed populations. This is particularly critical in the Arabian/Persian Gulf, where such events are frequent. The region's scarce freshwater sources necessitate reliance on seawater desalination, making the establishment of an early warning system for algal blooms vital for the efficient management of desalination plants and other economic activities.

In this master thesis, we explore the performance of a multi-sensor optical approach combining Sentinel-2 and Sentinel-3 images to detect algal blooms. For the first time, the study integrates the OLCI Terrestrial Chlorophyll Index (OTCI) to Sentinel-3 for algal bloom detection, and uses the Algae Bloom Detection Index (ABDI) with near-infrared (NIR) data from Sentinel-2. Through meticulous calibration, we simulated conditions for the year 2023, identifying risk zones across the Gulf with enhanced precision on specific dates when possible.

Our findings demonstrate the efficacy of using OTCI and Sentinel-3 for daily detection of algal blooms and validate the use of the combination of ABDI and NIR for precision detection with Sentinel-2, highlighting well algal blooms convergent fronts. However, the method primarily highlights the area on the water surface, potentially underestimating the blooms' total extent. While further refinement is needed to enhance the model's spatial and temporal accuracy, this research represents a decisive first step toward developing an early detection system. Despite its limitations, the study holds promise for achieving more comprehensive and large-scale application in the future.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
**Faculté des bioingénieurs**
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-La-Neuve, Belgique | www.uclouvain.be/agro