**AMY ZAMMIT**

# A COMPARATIVE STUDY OF MULTISPECTRAL AND HYPERSPECTRAL REMOTE SENSING TECHNIQUES FOR THE DETECTION OF CHLOROPHYLL-A CONCENTRATIONS IN COASTAL AREAS

**UNIVERSIDADE DO ALGARVE**
Faculdade de Ciências e Tecnologia
2024

**AMY ZAMMIT**

# A COMPARATIVE STUDY OF MULTISPECTRAL AND HYPERSPECTRAL REMOTE SENSING TECHNIQUES FOR THE DETECTION OF CHLOROPHYLL-A CONCENTRATIONS IN COASTAL AREAS

**Master’s in Marine and Coastal Systems**

**Performed under the supervision of:**
**Chaitanya Jalluri (Pixxel Space Technologies)**
**Joaquim Luis (UAlg)**

**UNIVERSIDADE DO ALGARVE**

Faculdade de Ciências e Tecnologia

2024

## Declaração de autoria de trabalho
## Declaration of Authorship of work

# A COMPARATIVE STUDY OF MULTISPECTRAL AND HYPERSPECTRAL REMOTE SENSING TECHNIQUES FOR THE DETECTION OF CHLOROPHYLL-*A* CONCENTRATIONS IN COASTAL AREAS

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

I declare to be the author of this work, which is original and unpublished. Authors and works consulted are duly cited in the text and are included in the list of references.

Amy Zammit
30th September 2024

## Abstract

Coastal water quality (CWQ) monitoring is essential for managing harmful algal blooms and assessing marine ecosystem health, with chlorophyll-*a* (chl-*a*) serving as a proxy for phytoplankton biomass. The study compared the Ocean Color x (Ocx) algorithm's performance in detecting chl-*a* concentrations using multispectral (Sentinel-2) and hyperspectral (Airborne Visible/InfraRed Imaging Spectrometer [AVIRIS]) imagery at three coastal sites, compared against *in-situ* measurements. Sentinel-2 provided moderate accuracy but struggled to capture small-scale variability in complex environments. Conversely, AVIRIS showed better detection due to its finer spectral resolution, though its accuracy was inconsistent, likely due to data processing limitations and environmental factors. The study suggests the integration of both imaging techniques to enhance CWQ monitoring and highlights the need for refined algorithms and aligning *in-situ* data collection with remote sensing acquisitions.

# Resumo

Esta dissertação investiga a eficácia das tecnologias de deteção remota multi e hiperespectral na deteção e quantificação das concentrações de clorofila-a (chl-a) em três ambientes costeiros distintos: o Canal de Santa Barbara (SBC), a Baía de Newport (NB) e a Lagoa de Veneza (LoV). A clorofila-a é um indicador chave da saúde dos ecossistemas aquáticos, utilizado para monitorizar a biomassa de fitoplâncton e detectar o aparecimento de algas nocivas (HABs), que afetam a vida marinha, a saúde humana e o ambiente. Dado o aumento global da frequência dos HABs, técnicas eficazes e escaláveis para sua detecção precoce são cruciais. Embora os métodos tradicionais, como a amostragem in-situ, sejam valiosos, eles são limitados na sua cobertura espacial e são dispendiosos. Em contraste, a deteção remota oferece uma monitorização contínua em grande escala de parâmetros como as concentrações de chl-a.

As tecnologias de deteção remota, como o Sentinel-2 (multiespectral) e o AVIRIS (hiperespectral), têm sido cada vez mais utilizadas para monitorizar águas costeiras. Os sensores multiespectrais adquirem dados num número limitado de bandas espectrais largas, adequados para identificar tendências gerais, mas limitados na deteção de mudanças subtis em parâmetros específicos, como a chl-a. Já os sensores hiperespectrais, com centenas de bandas espectrais estreitas, têm uma maior resolução espectral, permitindo uma deteção mais precisa das variações da chl-a e de outros constituintes da água.

Este estudo teve como objetivo avaliar e comparar o desempenho do Sentinel-2 e do AVIRIS na deteção de chl-a em três locais distintos, com diferentes desafios para a deteção remota. O Canal de Santa Barbara apresenta um ecossistema complexo, com influências marinhas e terrestres. A Baía de Newport é um ambiente mais urbanizado, com influências antropogénicas e diferentes graus de poluição. A Lagoa de Veneza é uma área bio-diversa e sensível, sujeita a alterações induzidas por atividades humanas. Esses locais são ideais para testar tecnologias de deteção remota em ambientes costeiros variados.

A metodologia envolveu a aquisição de imagens do Sentinel-2 e do AVIRIS para cada local de estudo, complementadas por dados de referência de medições de chl-a. As imagens do Sentinel-2 foram escolhidas pela sua resolução espacial e capacidade de fornecer dados em intervalos regulares. O AVIRIS, devido à sua alta resolução espectral, foi selecionado pela capacidade de detectar variações subtis nas concentrações de chl-a. Os dados foram pré-processados para

garantir compatibilidade com os modelos analíticos e comparados com as medições in-situ para avaliar a precisão das previsões baseadas em deteção remota.

Os resultados mostraram diferenças significativas entre o desempenho dos sensores. O Sentinel-2 teve um desempenho moderado no SBC, com erros relativamente baixos em comparação com os dados de referência, mas teve dificuldades na Baía de Newport e na Lagoa de Veneza, onde a resolução espectral mais ampla do sensor não conseguiu capturar variações de pequena escala nas concentrações de chl-a. A complexidade das águas costeiras e a presença de partículas em suspensão e substâncias dissolvidas também contribuíram para esse desempenho limitado.

Em contraste, o AVIRIS foi mais eficaz na deteção de chl-a, especialmente na Baía de Newport e na Lagoa de Veneza, onde sua resolução espectral fina permitiu capturar variações menores e distinguir melhor a chl-a de outros componentes da água. No entanto, no Canal de Santa Barbara, o desempenho do AVIRIS foi afetado por erros potenciais no pré-processamento e pela complexidade ambiental.

A pesquisa também destacou a importância de alinhar os dados de referência com as imagens remotas. Desajustes temporais entre as passagens dos satélites e a coleta de dados in-situ podem ter causado discrepâncias nos resultados. Além disso, foi identificado que refinamentos nos coeficientes dos algoritmos OCx podem melhorar a precisão das previsões de chl-a.

Em conclusão, embora tanto os sensores multiespectrais quanto os hiperespectrais tenham potencial para monitorizar as concentrações de chl-a, os resultados indicam que os sensores hiperespectrais, como o AVIRIS, oferecem vantagens significativas em ambientes costeiros, especialmente em áreas onde a variação de pequena escala é importante. Este estudo também ressaltou a necessidade de melhorias contínuas nos algoritmos de deteção remota e na validação de dados de referência, o que pode aprimorar a precisão e a aplicabilidade dessas tecnologias para monitorização ambiental em grande escala. O avanço destas tecnologias, em particular os sensores hiperespectrais, terá um papel crucial na mitigação dos impactos da poluição da água e na preservação dos ecossistemas costeiros.

# Table of Contents

Table of Figures

Table of Tables

Abbreviations

| | |
|---|---|
| .csv | comma-separated value |
| AAOT | Acqua Alta Oceanographic Tower |
| AOI | Area of interest |
| ASP | Amnesic Shellfish Poisoning |
| AVIRIS | Airborne Visible/InfraRed Imaging Spectrometer |
| Chl-*a* | Chlorophyll-*a* |
| CWQ | Coastal water quality |
| CZCS | Coastal Zone Color Scanner |
| EEA | European Environment Agency |
| EPA | Environmental Protection Agency |
| EPA | Environmental Protection Agency (USA) |
| ESA | European Space Agency |
| EU | European Union |
| HAB | Harmful algal bloom |
| HSI | Hyperspectral imaging |
| IOP | Inherent optical properties |
| L1 | Level-1 |
| LoV | Lagoon of Venice |
| MAE | Mean Absolute Error |
| MODIS | MODerate Resolution Imaging Spectroradiometer |
| MSI | Multispectral imaging |
| NB | Newport Bay |
| NIR | Near InfraRed |
| NRT | Near real-time |
| OCx | Ocean Color x |
| PnB | Plumes and Blooms (program) |
| RGB | Red, green, blue |
| RMSE | Root Mean Squared Error |
| SBC | Santa Barbara Channel |
| SeaBASS | SeaWiFS Bio-optical Archive and Storage System |
| SeaWiFS | Sea-Viewing Wide Field-of-View Sensor |

| | |
|---|---|
| SNAP | Sentinel Application Platform |
| SNR | Signal-to-noise ratio |
| SWIR | Short Wave InfraRed |
| TOA | Top-of-atmosphere |
| TSM | Total suspended matter |
| USC_SEAPRISM | University of Southern California SeaPRISM |
| VNIR | Visible and near-infrared |
| WHO | World Health Organization |

# 1. Introduction

Coastal water quality (CWQ) monitoring is essential for understanding the health and productivity of marine ecosystems. Coastal environments are highly dynamic and often face pressures from human activities. Therefore, regular assessment is necessary to mitigate potential risks to the ecosystem's balance (Abu Daya, 2004). Among the various CWQ parameters, chlorophyll-*a* (chl-*a*) is a direct proxy for phytoplankton biomass, indicating the state of primary producers. It is usually related to excessive nutrient input, potentially inducing harmful algal blooms (HABs), thus causing significant ecological, economic, and public health risks. Hence, early detection and monitoring of chl-*a* concentrations represent a crucial prerequisite of integrated environmental management (Gregor & Maršálek, 2004; Silva et al., 2016).

While traditional methods, such as *in-situ* sampling, are highly accurate in measuring chl-*a*, they are labour-intensive, time-consuming, and limited in spatial and temporal coverage (Duan et al., 2013a; Duan et al., 2013b). However, remote sensing technologies have led the shift towards more efficient and expansive monitoring. Remote sensing offers a comprehensive view of vast areas, providing frequent, near real-time (NRT) data, which is particularly useful for dynamic coastal environments (Gholizadeh et al., 2016).

Multispectral imaging (MSI) sensors, equipped on airborne or space-borne platforms, have been extensively used to monitor CWQ because of their high spatial resolution and frequent revisit times. However, multispectral sensors are limited by their broad spectral bands, which may not capture the subtle variations in chl-*a* concentrations, especially in optically complex coastal waters. Hyperspectral imaging (HSI), on the other hand, provides higher spectral resolution and enables more detailed analysis, including the differentiation between various water constituents (Ignaciov, 2022).

This dissertation aims to compare the performance of multispectral and hyperspectral imagery in detecting and quantifying chl-*a* concentrations in coastal environments. By focusing on three study sites—the Santa Barbara Channel (SBC), Newport Bay (NB), and the Lagoon of Venice (LoV)—the study assesses the accuracy and applicability of Sentinel-2 and Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) imagery in chl-*a* detection. The specific questions were addressed in the study:

1. What is the accuracy and effectiveness of Sentinel-2 (MSI) and AVIRIS (HSI) in detecting chl-*a* in coastal environments? Specifically, does the significantly higher spectral resolution provide better detection capabilities than MSI?
2. What are the limitations of multispectral and hyperspectral imagery in CWQ monitoring?
3. Could HSI contribute to developing future methodologies to create early warning systems for HABs?

Considering these questions, the dissertation aims to contribute to the emerging research on remote sensing for CWQ monitoring.

## 2. Literature Review

The following section briefly overviews remote sensing for chl-*a* monitoring in coastal waters. Beginning with a definition of CWQ parameters (Section 2.1) and moving on to chl-*a* and HABs (Sections 2.1.1 and 2.1.2), this section explores the role of remote sensing technology in coastal monitoring. It then delves into the advantages of remote sensing over traditional methods (Section 2.2.1) and examines the role of satellite-based imaging in detecting and monitoring HABs (Section 2.2.2). Additionally, it explores the rise of HSI (Section 2.3) compared to MSI, focusing on a notable program: AVIRIS (Section 2.3.2).

### 2.1. Coastal water quality (CWQ) parameters

Coastal waters are defined as surface “waters located outside the low-water line or the outer limit of an estuary”, extending one nautical mile seaward (European Environment Agency [EEA], 1991). In general, CWQ (commonly interchanged with “bathing water quality”) refers to the microbiological and physiochemical properties of surface waters (EEA, 2024). Various water quality parameters are critical for maintaining ecological balance and supporting marine life, which influence the health of coastal waters. Ritchie and Schiebe (1998) listed suspended sediments, chlorophyll, chemical substances, dissolved organic matter, nutrients, pesticides, thermal releases, and oils as water quality components.

Studies used measurable components, particularly chl-*a,* to assess aquatic ecosystems’ status and support economic sectors, including tourism, shipping, and fisheries (Gorgoglione et al., 2016; Bian et al., 2015; Grashorn et al., 2015). This section explores the significance of chl-*a* in determining the primary productivity of water bodies and, ultimately, supporting the early detection of eutrophication and HAB events.

### 2.1.1. Chlorophyll-*a* (chl-*a*)

Chl-*a*, a primary pigment found in all photosynthetic organisms, is an indicator of primary productivity and total phytoplankton biomass in aquatic environments (Gregor & Maršálek, 2004). Phytoplankton is comprised of various taxonomic groups such as chlorophytes, cryptophytes, cyanobacteria, diatoms, and dinoflagellates that form the base of the marine food web, influencing ecosystem processes (Lunetta et al., 2009; Paerl et al., 2003; Pinckney et al., 1998). Therefore, chl-*a* is correlated to nutrient concentration and algal production, providing insights into habitats' ecological status and composition (Gholizadeh et al., 2016; Moses et al., 2012; Moses et al., 2014).

High chl-*a* concentrations often indicate nutrient-rich conditions, which can lead to eutrophication and subsequent ecological imbalances. Gruber and Galloway (2008) found waters near densely populated areas to experience environmental stresses from excessive nutrient input. Besides causing aesthetic problems, including odours and murkiness, such conditions can cause anoxic conditions and have devastating effects on sensitive ecosystems (World Health Organization [WHO], 2006, p. 166; Environmental Protection Agency [EPA], 2024). Thus, efficient monitoring of chl-*a* provides an in-depth understanding of eutrophication frequency and harmful effects on bathing waters' environmental and socioeconomic status (Louis et al., 2005).

Monitoring chl-*a* concentrations is commonly done through *in-situ* methods aboard research vessels or buoys and, more recently, remote sensing techniques (Section 2.2.1). *In-situ* determination of chl-*a* involves collecting, filtering, and storing water samples before using a spectrometer to determine absorption characteristics at specific wavelengths. Although these measurements are highly accurate, they are labour-intensive and time-consuming. Therefore, it is not economically feasible to regularly monitor chl-*a* in coastal areas (Duan et al., 2013a; Duan et al., 2013b). Therefore, remote sensing of CWQ parameters using satellites is an attractive option (Gholizadeh et al., 2016).

### 2.1.2. Harmful Algal Blooms (HABs)

HABs are a significant environmental concern in coastal areas, with increasing frequency and geographical spread worldwide (Silva et al., 2016). These blooms, which promote toxinproducing phytoplankton growth, can have severe ecological, economic, and public health impacts, including oxygen depletion, massive fish kills, and drinking water contamination. HABs produce various biotoxins that induce syndromes such as paralytic, diarrhetic, and amnesic shellfish poisoning, with most toxins synthesised by dinoflagellates (Anderson, 2009). Notably, domoic acid, responsible for Amnesic Shellfish Poisoning (ASP), is produced by diatoms (Silva et al., 2016). These events severely restrict the sustainable development of coastal regions and endanger public health (Trainer et al., 2010). As reported by monitoring programs, the main species contributing to HABs include *Pseudo-nitzschia*, *Dinophysis*, *Gymnodinium*, and, more recently, *Ostreopsis* and *Karenia*. These species exhibit distinct oceanographic preferences and adaptive strategies, reflecting diverse life-form characteristics (Moita et al., 2006; Palma et al., 2010; Silva et al., 2009; Smayda, 2002).

Monitoring and early detection of HABs are crucial for protecting marine ecosystems and human health. Current monitoring methods rely on diverse approaches, including field observations and advanced computational models. HAB alert systems utilise a variety of data sources, including ocean colour satellite imagery, models, historical trends, public health reports, buoy data, and forecasts of bloom progression (Anderson et al., 2015).

Remote sensing technologies can monitor chl-*a* concentrations over large spatial scales, providing valuable information for early HAB detection and management (Lock et al., 2023). By integrating remote sensing data with traditional methods, researchers can gain a comprehensive understanding of CWQ dynamics and effectively mitigate the impacts of HABs on marine ecosystems and human communities.

### 2.2. Remote sensing to monitor and detect CWQ parameters

Remote sensing technologies have revolutionised the monitoring of CWQ. Unlike traditional *in-situ* sampling methods, remote sensing enables the systematic and cost-effective assessment of water quality parameters over large spatial and temporal scales. Remote sensing provides NRT and long-term monitoring capabilities, which are crucial for climatology and trend identification. It enables a view that *in-situ* measurements alone cannot reproduce (Gholizadeh et al., 2016; Abu Daya, 2004).

Monitoring the Earth's oceans through remote sensing officially began in 1978 with the launch of the Coastal Zone Color Scanner (CZCS), the first ocean colour sensor aboard NASA's Nimbus-7 satellite ("Supported Missions," n.d.). The CZCS was designed to measure chlorophyll concentration, sediment content, and other properties in coastal and oceanic waters, providing critical insights into biological productivity and water quality. This mission laid the groundwork for subsequent ocean colour remote sensing programs, including the European Space Agency's (ESA) Envisat platform in 2002.

Over the last decade, remote sensing of CWQ has advanced with the launch of satellite missions like the European Union's (EU) Copernicus Sentinel-2 and Sentinel-3 programs (Copernicus, n.d.). These satellites carry optical sensors, including the Multispectral Instrument onboard Sentinel-2, which captures high-resolution data on various CWQ parameters.

Integrating satellite-based remote sensing with other technologies offers comprehensive insights into CWQ dynamics. The following section explores the advantages of remote sensing over traditional methods and discusses its application in detecting increased chlorophyll concentrations and resultant HABs.

### 2.2.1. Why use remote sensing over traditional methods?

*In-situ* methods, while highly accurate, are often labour-intensive, time-consuming, and limited in spatial coverage. In contrast, remote sensing enables continuous, large-scale, and NRT monitoring of water quality parameters, enabling the detection of spatial and temporal patterns that would be missed by *in-situ* methods alone. Remote sensing captures vast areas, providing comprehensive insights into broader ecological dynamics and trends. Various methods exist for monitoring algal concentrations and detecting HABs, including molecular biology approaches, large-scale field surveys, numerical modelling, and remote sensing through satellites (Silva et al., 2016). While *in-situ* species-specific cell counts provide highly accurate data, their spatial resolution is limited.

However, despite its advantages over *in-situ* measurement techniques, remote sensing has notable limitations. Sensors cannot penetrate cloud data, which is particularly problematic during storms or in high latitudes, rendering images unavailable for further analysis (Holt et al., 2017). Additionally, remote sensing is inherently limited to surface-level observations, as satellite sensors can only detect and measure parameters at or near the water's surface. This limits coastal water analysis, where subsurface dynamics such as stratification or deep-water nutrient mixing significantly impact water quality but remain undetectable by remote sensing instruments.

The availability of imagery is constrained by the fixed overpass times of polar-orbiting (specifically sun-synchronous) satellites, meaning end-users cannot capture data at a time of their choosing but must rely on the satellite's predetermined orbit (ESA, 2020). These limitations restrict the data available to end-users, especially when analysing a single phenomenon, like an extreme weather event or HAB.

Operational models enable NRT forecasting, facilitating the calculation of phytoplankton distribution and chlorophyll concentration (Silva et al., 2016; Stumpf et al., 2010). However, the accuracy of retrieved physical, chemical, or biological parameters depends on trained and calibrated models. Monitoring is only as accurate as the procedures and algorithms (Gholizadeh et al., 2016). Algorithms designed for land or open water monitoring tend to face issues when computing parameters in optically complex and turbid coastal waters (Abu Daya, 2004; Moses et al., 2012).

The dynamic nature of coastal waters requires a combination of *in-situ* and high spatial, spectral, and temporal resolution data from remote sensing to overcome the limitations of both techniques. Efficient monitoring involves collecting and analysing remotely sensed data due to its accessibility, low cost, and continuous nature and validating findings through *in-situ* observations (Nazeer et al., 2017; Gholizadeh et al., 2016). National and international legislation, such as the EU’s Water Framework Directive, advise this method for coastal water monitoring (Directive 2000/60/EC).

### 2.2.2. Detecting HABs using satellite-based remote sensing

Satellite-based remote sensing has become a powerful tool for monitoring and detecting HABs in coastal waters. Sensors like the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) have extensively monitored chl-*a* concentrations to identify HABs. These sensors provide data at different spatial and temporal resolutions, enabling researchers to track the development and movement of blooms over time. Remote sensing, particularly satellite-based imaging, revolutionises the assessment of spatial and temporal variability in CWQ, overcoming limitations in data availability from remote or large marine areas. These advancements have ushered in diverse applications in aquaculture, hazard detection, and environmental management (Fernández-Tejedor et al., 2022). However, retrieving parameters such as chl-*a* concentration remains challenging, primarily due to the complex optical properties and high spatial variability of coastal regions (Fernández-Tejedor et al., 2022; Lunetta et al., 2009). Therefore, atmospheric correction algorithms are crucial for accurate analysis, particularly in Case-2 waters (i.e., coastal and inland waters) where the physical (e.g., depth) and biochemical (e.g., organic matter, salinity, etc.) characteristics alter the optical properties (Fernández-Tejedor et al., 2022; Lock et al., 2023).

Remote sensing techniques allow for monitoring of the extent and intensity of HAB events. Optical sensors detect changes in reflected solar spectra caused by variations in CWQ parameters, such as chlorophyll concentration, turbidity, and coloured dissolved organic matter (CDOM). These spectral variations provide crucial information for estimating the concentration of these parameters in water, making optical sensors crucial for tracking and understanding coastal and marine (Ritchie et al., 2003; Abu Daya, 2004). The Sentinel-2 MSI platform, for instance, measures Earth's reflected radiance from Near InfraRed (NIR) to Short Wave InfraRed (SWIR) imagers in 13 spectral bands, including the visible spectra (i.e., red, green, blue [RGB]). Level-1 (L1) products refer to top-of-atmosphere (TOA) radiances in sensor geometry (L1B) and cartographic geometry (L1C). L1C processing elaborates on L1B products using resampling algorithms and cloud and land/water mask computations (ESA, n.d.-a).

Chl-*a* appears green as it absorbs energy from violet-blue and orange-red light. As concentrations increase, spectral responses at short wavelengths decrease, particularly in the blue band, enabling their detection by satellite sensors (Gholizadeh et al., 2016). Numerous algorithms exist to generate chl-*a* concentrations from remotely sensed imagery. Bio-optical models based on the inherent optical properties (IOP) of waterbodies (such as the absorption and scattering coefficients) and expected concentration ranges determine concentrations through inversion of water-leaving reflectance and atmospheric correction (Brockmann et al., 2016). Chl-*a*'s ability to be retrieved by satellite optical sensors across various spectral bands highlights its potential as a use case for the applications of HSI.

### 2.3. The rise of hyperspectral imaging (HSI)

In recent years, HSI has emerged as a revolutionary technology for environmental monitoring, offering unprecedented capabilities for spectral analysis and detection. This section explores the advantages of HSI over MSI for assessing CWQ parameters, focusing on its ability to provide detailed spectral information for accurate and comprehensive analysis. Additionally, the notable HSI program, AVIRIS, is discussed to illustrate the potential of this technology in advancing our understanding of coastal ecosystems.

### 2.3.1. Why use hyperspectral imaging over multispectral imaging?

While MSI sensors capture data in a limited number of broad wavelength bands, hyperspectral sensors collect data across hundreds of narrow, contiguous spectral bands (Figure 1). The high spectral resolution enables the precise identification and quantification of various substances in the water, including chl-*a*. HSI can detect subtle spectral features that are not detectable with MSI sensors, resulting in more accurate water quality assessments (Ignaciov, 2022). For instance, hyperspectral sensors mounted on small aircraft have been used for various water quality purposes, such as determining aquatic health status, supporting the identification of phytoplankton species, and assessing their spatial and temporal distribution (Olmanson et al., 2012).

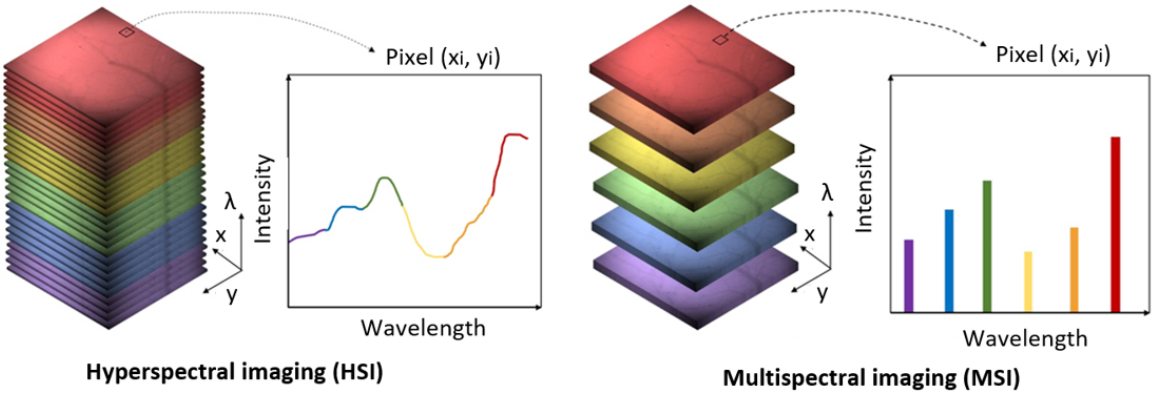

Figure 1: The difference between HSI and MSI (Ignaciov, 2022).

However, despite its advantages, HSI's voluminous and complex data requires significant computational specifications for processing and analysis (Vandermeulen et al., 2020). Additionally, the cost and availability of hyperspectral sensors, especially for large-scale applications, often present a challenge (Dekker et al., 2018). While HSI offers significantly higher spectral resolution, it is often limited in spatial resolution, narrower swaths, and less frequent revisit times, making it difficult to track subtle changes over time. However, recent advancements due to smaller sensors, reduced costs, and the emergence of small satellites (small-sats) have led to the commercialisation of HSI satellites that improve the technological specifications of the sensors and products while also making imagery accessible to a broader range of industries and organisations (Erwin, 2023).

The ability of hyperspectral sensors to provide continuous spectral coverage throughout the visible and near-infrared (VNIR) spectrum allows for detailed spectral sampling that can track subtle differences in surface waters over time (Dierssen et al., 2021). The high spectral resolution of HSI systems, with bandwidths typically less than 15 nm and more than 20 bands in visible wavelengths, enables the development of sophisticated algorithms to analyse water quality parameters (Dierssen et al., 2021). Furthermore, studies have shown that hyperspectral sensors with bandwidths of 5 nm can provide even more detailed and accurate data for a wide range of aquatic environments (Wolanin et al., 2016; Vandermeulen et al., 2017; Dekker et al., 2018).

This capability to detect subtle variations in spectral signatures is crucial for understanding and managing coastal ecosystems and monitoring the impacts of environmental stressors such as pollution and climate change (Lunetta et al., 2009). Additionally, HSI can support new algorithm concepts that require continuous and dense spectral sampling, further enhancing its potential applications in environmental monitoring (Dierssen et al., 2021).

### 2.3.2. Notable programs: A highlight on AVIRIS

Operated by NASA, AVIRIS is one of the most notable HSI programs supporting the monitoring and analysis of CWQ. AVIRIS collects data between 400 and 2500 nm, providing sufficient spectral information to analyse various environmental parameters, including chl-*a* concentrations. Its high spatial and spectral resolution–measuring radiance spectra as images of 11 km width and up to 800 km length with a spatial resolution between 4 and 20 meters, depending on flight altitude–makes it a valuable tool for monitoring coastal ecosystems, detecting HABs, and conducting comprehensive environmental assessments (Green et al., 1998).

AVIRIS data have been applied in numerous studies, demonstrating the potential of HSI for CWQ monitoring. For example, AVIRIS has been used to estimate surface concentrations of chl-*a*, CDOM, and suspended solids in various water bodies (Lunetta et al., 2009). The sensor's ability to generate detailed spectral data without needing a priori band selection allows for the development of sophisticated algorithms that enhance the accuracy of water quality assessments (Hoogenboom et al., 1998). The high signal-to-noise ratio (SNR) of AVIRIS also supports researchers in conducting detailed analyses of coastal and estuarine environments.

One significant application of AVIRIS is its role in estimating chl-*a* concentrations using spectral band ratios. By detecting the chl-*a* absorption maximum at 673.6 nm and correlating it with measured water sample concentrations, researchers can generate accurate estimates of chl-*a* distribution across study areas (Lunetta et al., 2009).

Additionally, AVIRIS data have been used to monitor the emergence and extent of intense phytoplankton blooms, including HABs. For instance, AVIRIS imagery has been employed to detect and quantify phytoplankton blooms in coastal regions, providing valuable information to manage these events and mitigate their impacts on marine ecosystems and human health (Palacios et al., 2015).

Moreover, integrating hyperspectral imagery with other remote sensing data and environmental parameters can address a wide range of societal problems, including the response of coastal ocean ecosystems to population growth and climate change (Muller-Karger et al., 2018). Hyperspectral data provide potential new applications that are not limited by gaps in the spectrum. For example, the ability to differentiate between various water masses and track subtle differences over time can enhance our understanding of coastal ecosystem dynamics and support the development of effective management strategies (Dierssen et al., 2021).

In conclusion, MSI and HSI are both vital for detecting and monitoring chl-*a* concentrations and other CWQ parameters. While MSI provides broad and continuous monitoring capabilities, HSI offers unparalleled spectral resolution for detailed analysis. Integrating these technologies enhances our ability to understand and manage coastal ecosystems effectively. Hyperspectral remote sensing, with its potential to provide new applications relevant to societal needs, such as assessing aquatic biodiversity and habitats, is poised to play a critical role in future environmental monitoring and management strategies (Dierssen et al., 2021).

## 3. Methodology

This section outlines the methods used to detect and quantify chl-*a* concentrations in coastal environments. It covers the study site selection (3.1), providing information on the areas of interest (AOIs). It also outlines how Sentinel-2 and AVIRIS imagery and ground-truth data were collected and processed (3.2). Lastly, it also explains the statistical and comparative techniques employed to compare the imagery-derived results against *in-situ* measurements (3.3). Limitations encountered are also discussed (3.4).

### 3.1. Study sites

The selected study sites experience high primary productivity and susceptibility to algal blooms. These locations include the SBC in California (3.1.1), a dynamic coastal ecosystem influenced by local wind patterns and large-scale ocean currents; NB in California (3.1.2), an estuarine environment impacted by nutrient loading and urban runoff; and the LoV in Italy (3.1.3), a shallow lagoon facing eutrophication pressures due to human activities.

### 3.1.1. The Santa Barbara Channel (SBC), California, North America

The SBC (Figure 2) is situated between the California coast and the Channel Islands chain at the northern edge of the Southern California Bight. The basin, approximately 100 km by 40 km, features complex bathymetry with a deep central basin and narrow continental shelves along its northern and southern boundaries. It is separated from the Southern California Bight by a 200 m deep sill to the east and a 400 m deep sill to the west (Hendershott & Winant, 1996).

Figure 2: Study area showing the imagery tiles downloaded, including multi- (blue) and hyperspectral (red) imagery and the ground-truth sampling stations (yellow) for SBC.

The region experiences mixing from warm, saline waters from the Southern California Bight and colder, fresher waters from upwelling near Pt. Conception and Pt. Arguello, with seasonal upwelling intensifying during the spring due to equatorward winds (Hendershott & Winant, 1996; Hickey, 1979; Lynn & Simpson, 1987). The SBC supports diverse ecosystems, including kelp forests and intertidal habitats, and is a popular site to study primary productivity due to its higher phytoplankton biomass compared to the surrounding Bight (Beers et al., 1986; Levin et al., 2015). Its proximity to the Santa Clara River, the largest watershed draining into the channel, also brings significant terrestrial nutrient input (Mertes & Warrick, 2001; Otero & Siegel, 2004).

Algal blooms and high chlorophyll concentrations in the SBC are driven by nutrient-rich upwelling waters. Seasonal upwelling, enhanced by local wind stress curl, brings cold, nutrient-dense water to the surface, forming ideal conditions for phytoplankton growth (Jones et al., 1988; Venrick, 1998; Dorman & Winant, 2000; Otero & Siegel, 2004). The channel's relatively sheltered environment and complex bathymetry help retain these nutrients (Mertes & Warrick, 2001). Consequently, the timing and intensity of blooms are closely tied to these physical and oceanographic processes (Eppley et al., 1979; Otero & Siegel, 2004).

### 3.1.2. Newport Bay (NB), California, North America

NB (Figure 3) is Southern California's second-largest estuarine embayment, supporting habitats for aquatic and terrestrial species. San Diego Creek is the primary freshwater input, draining approximately 85% of the surrounding 400 km² urbanised watershed. Therefore, the bay is subject to significant anthropogenic stressors, including nutrient loading from agricultural runoff and urban development, causing eutrophication and macroalgal blooms (Kamer et al., 2001; Nezlin et al., 2007).

Water circulation in the upper bay is influenced by tidal patterns, with residence times averaging about one week during neap tides (Kennison et al., 2003). The lower bay, developed into a marina, has diminished much of the natural wetland habitat. In contrast, the upper bay retains more ecological characteristics but is considered eutrophic due to nutrient influx from both point and non-point sources (Nezlin et al., 2009).

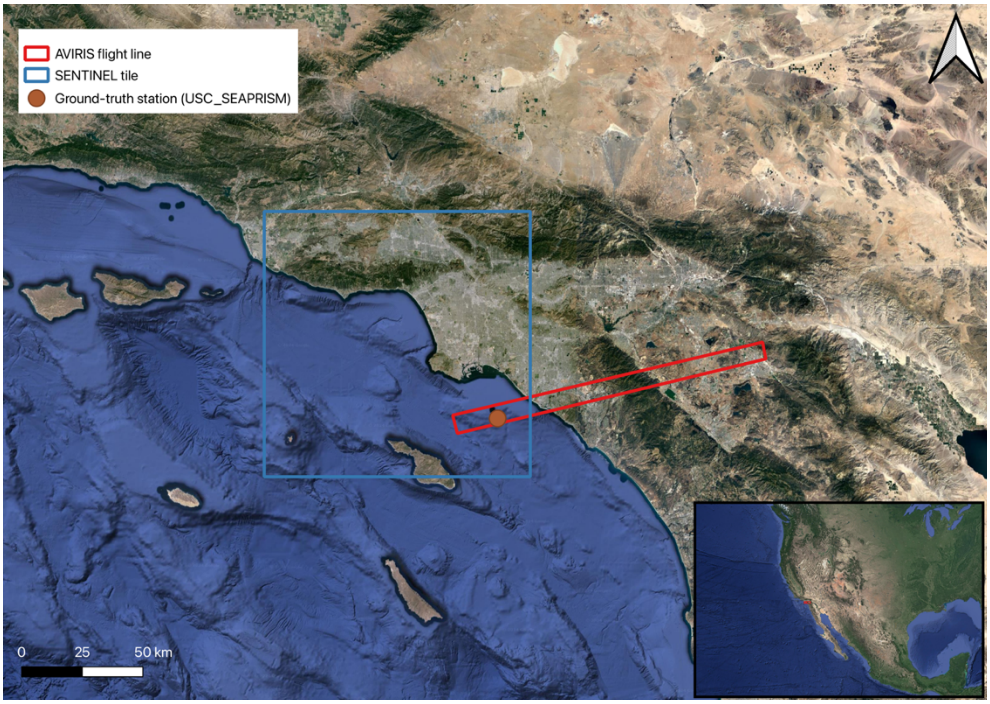

Figure 3: Study area showing the imagery tiles downloaded, including multi- (blue) and hyperspectral (red) imagery, and the ground-truth sampling station (brown) for NB.

Persistent macroalgal blooms, dominated by species such as *Ulva* spp. and *Ceramium* spp., lead to hypoxic conditions during decomposition, particularly in poorly flushed areas and warmer months (Nezlin et al., 2007). Studies found that hypoxia has been a recurring issue since the 1970s, with macroalgal growth significantly disrupting oxygen dynamics, particularly in areas with limited vertical stratification due to shallow depths and tidal mixing (Kennison et al., 2003). Efforts to mitigate nutrient inputs, such as the 1998 Total Maximum Daily Load for nitrogen and phosphorus, are ongoing (Kennison et al., 2003). However, eutrophication remains a concern, requiring continued monitoring of water quality in the bay.

### 3.1.3. Lagoon of Venice (LoV), Italy, Europe

The LoV (Figure 4), located in the northern Adriatic Sea, spans approximately 550 km², making it Italy's largest lagoon. It has an average depth of around 1 metre, with deeper channels (10-15 metres) connecting to the sea through three inlets: Lido, Malamocco, and Chioggia (Aubry et al., 2021; Acri et al., 2020; Facca et al., 2014). Tidal water flows through these inlets, influencing the lagoon's hydrodynamics, with tidal excursions averaging around 31 cm but occasionally reaching up to 170 cm under certain meteorological conditions (Acri et al., 2020). Over centuries, human interventions have aimed to reduce siltation and protect the lagoon from the sea, facilitating its use as a commercial port and supporting industrial, urban, and agricultural development (Çevirgen et al., 2020). However, these interventions have disrupted the sediment balance, leading to significant erosion, particularly in the marshlands and intertidal areas. The lagoon loses approximately one million tons of sediment annually, deepening the shallow areas by over 50 cm in some regions (Aubry et al., 2021; Acri et al., 2020).

Figure 4: Study area showing the imagery tiles downloaded, including multi- (blue) and hyperspectral (red) imagery, and the ground-truth sampling station (green) for LoV.

The disruption of the lagoon’s sediment dynamics and the increased pollution have led to significant eutrophication and resultant HABs, dominated by opportunistic species. Despite efforts to reduce runoff, over 7000 tons of nitrogen and 1000 tons of phosphorus enter the lagoon annually from the drainage basin and urban-industrial areas (Acri et al., 2020).

In the 1970s, phytoplankton blooms were also present, but macroalgae like *Ulva rigida* now dominate, especially in nutrient-rich, shallow areas with limited water exchange (Çevirgen et al., 2020). The proliferation of *Ulva* is facilitated by several factors, including its ability to efficiently uptake and store excess nitrogen and thrive in the lagoon's environment. This shift has led to habitat loss, reduced biodiversity, and marshland erosion (Acri et al., 2020). The ongoing challenges make LoV a suitable site to assess chlorophyll concentrations, highlighting the importance of continued research and intervention to restore the lagoon's ecosystem and mitigate the impacts of nutrient pollution (Aubry et al., 2021; Poggioli, 2008).

### 3.2. Data acquisition

This section outlines the data collection performed for this study. Ground-truth data (3.2.1) were acquired from various sources to compare the imagery results later. Multispectral imagery from Sentinel-2 (3.2.2) was downloaded, processed, and analysed for the study areas, focusing on the relevant spectral bands for chlorophyll detection. Hyperspectral imagery from NASA's AVIRIS (3.2.3) sensor was also obtained and processed before being used for further analysis.

#### 3.2.1. Ground-truth data

Ground-truth data were collected from past missions that measured chl-*a* concentrations on the same dates and locations as the AVIRIS and Sentinel-2 orbit paths. For the SBC site (Figure 5), data were obtained from the Environmental Data Initiative repository. The dataset forms part of the 'Plumes and Blooms' (PnB) program (18th September 2019), which provides chl-*a* concentrations measured at various stations along the Santa Barbara coastline (SCB Marine Biodiversity Observation Network, Catlett, Siegel, Guillocheau, & Kui, 2024).

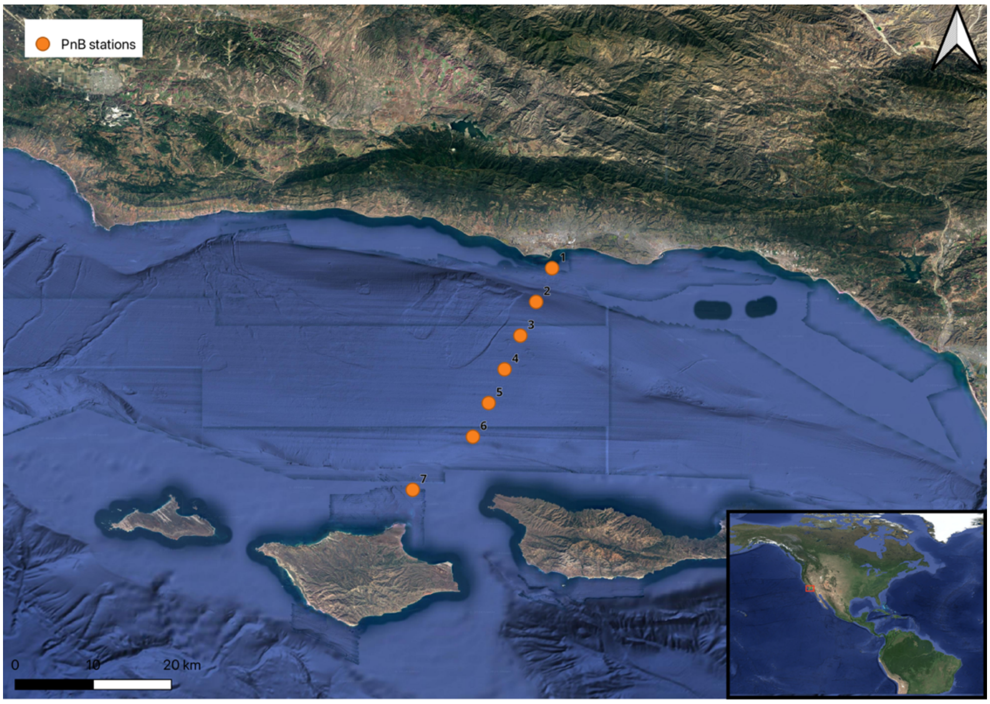

Figure 5: Locations of the PnB ground-truth stations within the SBC (coordinates in Appendix 1).

For the LoV and NB sites (Figure 6), data were obtained from NASA's AERONET-OC network, part of the SeaWiFS Bio-optical Archive and Storage System (SeaBASS). The Acqua Alta Oceanographic Tower (AAOT) was used for the LoV (4th June 2021), located eight nautical miles off the Venetian coast, while the University of Southern California “SeaPRISM" (USC_SEAPRISM) site (28th June 2023), located on an oil platform ten nautical miles offshore provided *in-situ* data for the NB area (AERONET, Zibordi, & Bulgarelli, 2024; AERONET, Jones, & Ragan, 2024). The coordinates and recorded chl-*a* values from these stations were organised into comma-separated value (.csv) files for later use.

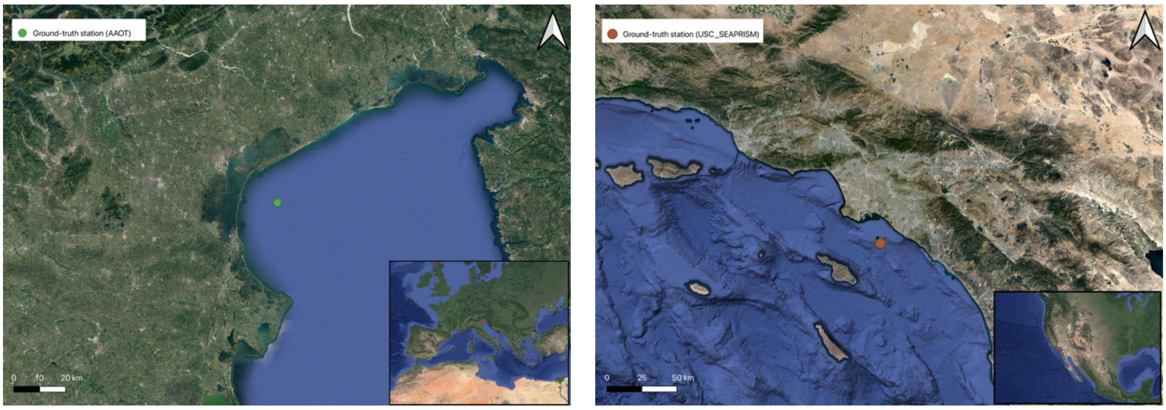

Figure 6: Location of the AAOT ground-truth station offshore from the LoV (45.314°N, 12.508°E) (left) and the USC_SEAPRISM ground-truth station offshore from NB (33.564° N, 118.118° W) (right).

### 3.2.1. Multispectral imagery

Sentinel-2 imagery was downloaded from the Copernicus Open Access Hub (https://browser.dataspace.copernicus.eu/). The Sentinel-2 mission, consisting of two polar-orbiting satellites (S2A and S2B), provides high spatial resolution imagery across 13 spectral bands, with spatial resolutions of 10 m, 20 m, and 60 m, depending on the spectral band. Sentinel-2 offers a revisit time of 5 days at the equator, making it particularly useful for monitoring dynamic coastal environments. The imagery was downloaded as Level-2A products, which include bottom-of-atmosphere reflectance data corrected for atmospheric effects.

The imagery selected corresponded to the study areas, with Sentinel-2 and AVIRIS imagery chosen based on the closest available dates to the ground-truth data collection (Table 1).. The data was downloaded as multi-band GeoTIFF files to support extracting and analysing individual spectral bands. Only bands relevant to chl-*a* analysis were retained, including Band 2 (Blue, 490 nm) and Band 3 (Green, 560 nm). The downloaded Sentinel-2 products were processed using the Sentinel Application Platform (SNAP), an open-source tool developed by the ESA (ESA, n.d.).

Table 1: Acquisition dates for ground-truth data readings and Sentinel-2 and AVIRIS imagery.

| Site | Sensor/Station | Date (dd/mm/yyyy) |
|---|---|---|
| SBC | PnB | 18/09/2019 |
| | Sentinel-2 | 19/09/2019 |
| | | 24/09/2019 |
| | AVIRIS | 25/09/2019 |
| NB | USC_SEAPRISM | 28/06/2023 |
| | Sentinel-2 | 27/06/2023 |
| | AVIRIS | 28/06/2023 |
| LoV | AAOT | 04/06/2021 |
| | Sentinel-2 | 04/06/2021 |
| | AVIRIS | 04/06/2021 |

After cropping the imagery to smaller AOIs (Figure 7) to facilitate processing, resampling was performed using Band 2 (10 m) to standardise the spatial resolution across the selected bands. A land mask was applied to ensure that chl-*a* calculations were limited to coastal waters (“California State Boundary,” 2018; “Municipal Boundaries of Italy 2019,” 2019).

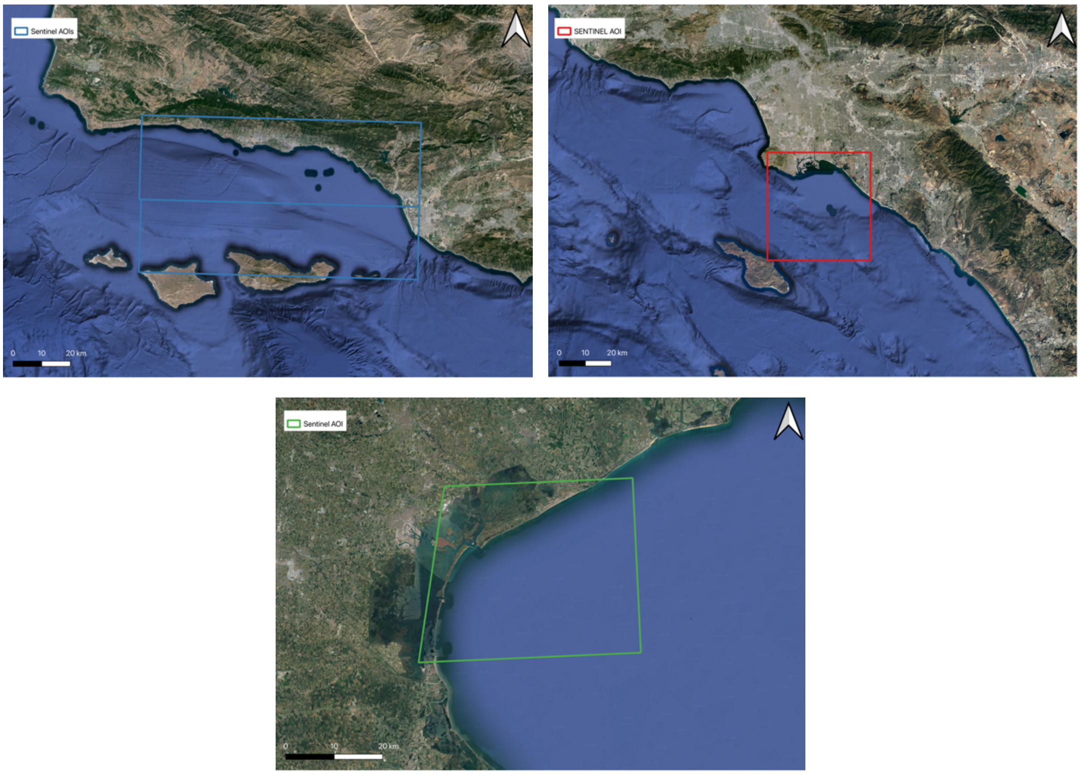

Figure 7: Cropped subsets for the SBC (top left), NB (top right), and LoV (bottom) were used to facilitate processing for the MSI analysis.

Chl-*a* concentration was calculated using the Ocean Color (OC2) algorithm, which involved applying a polynomial function to the reflectance ratio between the Blue and Green bands (Werdell et al., 2023). The coefficients ($a_0 = 0.3308$, $a_1 = -2.66837$, $a_2 = 1.59899$, $a_3 = 0.55248$, $a_4 = -1.4876$) were selected based on a study by Niroumand-Jadidi et al. (2021). The equation is defined as:

$$\log_{10}(chlor_a) = a_0 + \sum_{i=1}^{2} a_i \left(\log_{10}\left(\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})}\right)\right)^i$$

The logarithmic values were then converted to actual chl-*a* concentrations using:

$$chlor_a = 10^{(\log_{10}(chlor_a))}$$

The results were saved as GeoTIFF files (Appendix 1), maintaining the original spatial resolution and projection of the input imagery. The processed files were opened in QGIS for visualisation, where a logarithmic colour scale was applied to highlight chl-*a* concentration variations across the study areas. Lastly, ground-truth coordinates were loaded, and the 'Sample Raster Value' tool was used to extract and compare satellite-derived chl-*a* values with *in-situ* measurements.

### 3.2.2. Hyperspectral imagery

Hyperspectral imagery for the study areas was obtained from NASA's AVIRIS data portal (NASA Jet Propulsion Laboratory, n.d.-a). Only radiance data (Level-1B) were available for the SBC and NB, requiring preprocessing, while the LoV data were already available in reflectance format (Level-2A). Preprocessing followed the steps outlined in the ENVI Preprocessing Tutorial, using the FLAASH tool to convert radiance to reflectance, correcting for atmospheric effects, including absorption by water vapour, oxygen, carbon dioxide, methane, and ozone and molecular and aerosol scattering (ENVI Tutorials, n.d.). Scale factors from a preexisting gain file were applied (NASA Jet Propulsion Laboratory, n.d.-b). Sensor and atmospheric model parameters were set according to the characteristics of the AVIRIS sensor and the scene's environmental conditions (ENVI Tutorials, n.d.).

The preprocessing involved cropping the reflectance imagery to geographical subsets (Figure 8) and applying a land mask to reduce file size, facilitate processing, and focus the analysis on coastal waters.

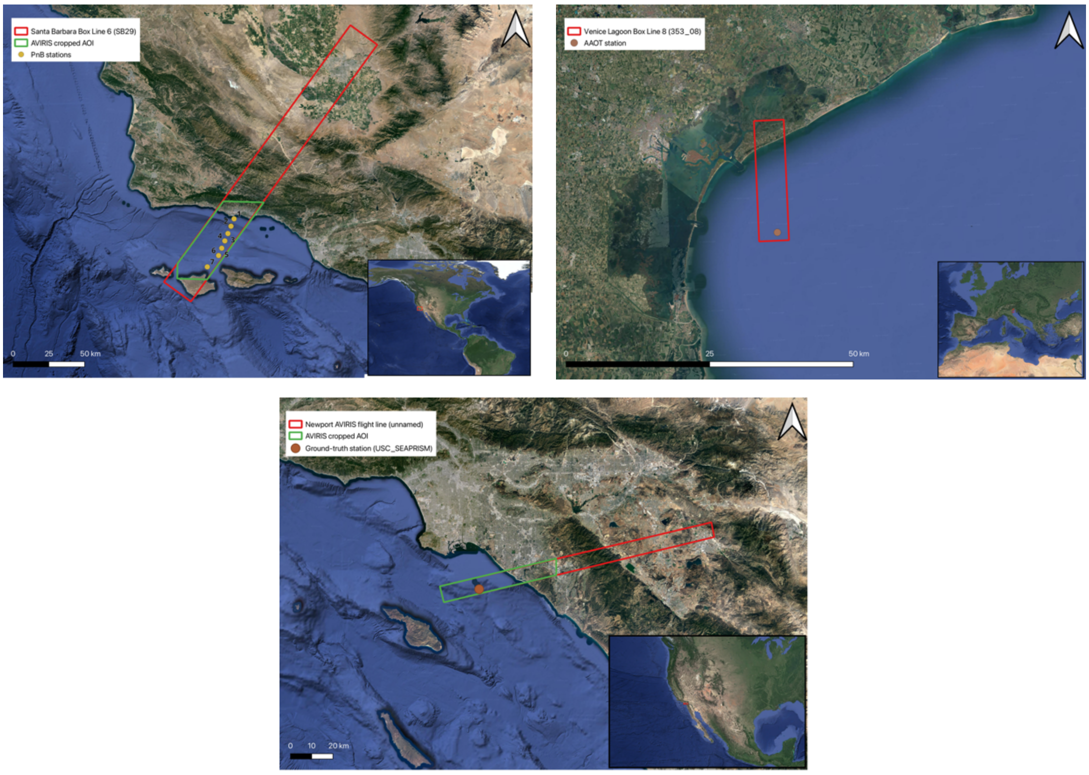

Figure 8: Study areas for the SBC (top left), NB (top right), and LoV (bottom) sites showing the downloaded hyperspectral (AVIRIS) imagery and ground-truth sampling stations. The maps for the SBC and NB include the AVIRIS flight lines (red) and cropped subsets to focus on coastal waters. The LoV imagery did not require a geographical subset due to its shorter flight path.

Reflectance imagery was used to estimate chl-$a$ concentrations using the OC3 algorithm (Section 3.2.1), with AVIRIS bands centred around 443 nm (Blue), 486 nm (Blue), and 551 nm (Green). In this case, a second blue band (486 nm) was used due to the high spectral fidelity of HSI, which provides the option to include multiple closely spaced bands. The chl-$a$ concentration calculation was based on a ratio between the maximum blue $R_{rs}$ value over a specified range (443 – 486 nm) and the green wavelength (551 nm). Since three $R_{rs}(\lambda)$ values are included, this equation is referred to as OC3 (O'Reilly & Werdell, 2019):

$$\log_{10}(chlor_a) = a_0 + \sum_{i=1}^{3} a_i \left( \log_{10} \left( \frac{\max[R_{rs}(\lambda_{blue})]}{R_{rs}(\lambda_{green}} \right) \right)^i$$

The algorithm utilised constants ($a_0 = 0.2521, a_1 = -2.2146, a_2 = 1.5193, a_3 = -0.7702$) adapted from Kolluru et al. (2023).

The processed data were saved as GeoTIFF files (Appendix 2) and visualised in QGIS, where the same logarithmic colour scheme was applied. Chl-$a$ values were extracted from the ground-truth station coordinates using the Sample Raster Values tool, enabling comparison of the chl-$a$ values derived from AVIRIS imagery against *in-situ* measurements.

### 3.3. Data analysis

The extracted chl-$a$ concentration data underwent statistical tests to compare against ground-truth readings and assess the accuracy of the sensors (Appendix 3). For SBC, where multiple ground-truth data points were available, several methods were applied to handle missing data due to cloud coverage or image swath limitations. These included: (1) dropping rows with NaN values, which reduces dataset size but ensures no missing values; (2) mean imputation, where NaN values are replaced with averages, retaining dataset size but potentially introducing bias; and (3) linear interpolation, which estimates values based on trends to preserve temporal and spatial continuity. Addressing NaNs is critical to ensure unbiased and accurate results.

For sites with only a single ground-truth data point (i.e., LoV and NB), the accuracy of satellite-derived chl-$a$ values was assessed by comparing them directly to the single available measurement. Given the limited data, no advanced statistical methods (such as MAE or RMSE) could be applied. Instead, the comparison provided a basic check to assess how closely the satellite-derived data matched the available ground-truth measurement.

For SBC, the normality of the data distribution was tested using the Shapiro-Wilk test to determine the appropriate statistical method for subsequent analysis. For normally distributed data, a paired t-test was used to compare the means of imagery-derived and ground-truth values. For non-normally distributed data, the Wilcoxon Signed-Rank test was used as a non-parametric alternative to evaluate the accuracy of the satellite imagery.

## 3.4. Limitations

Several limitations influenced the study and impacted both the methodology and results. A major challenge was the availability of datasets, which necessitated a significant alteration of the proposed plan. Instead of selecting a location and checking for available imagery and ground-truth data, the approach needed to be reversed; this meant identifying areas with available ground truth, AVIRIS, and Sentinel-2 data before finalising the study sites. This reversal proved problematic, as locating imagery from different sensors for the exact dates and locations as the ground-truth data was difficult.

Additionally, cloud coverage further constrained the selection of Sentinel-2 imagery. Significant clouds on September 24, 2019, obscured parts of the AOI and limited the extraction of chl-*a* values from several PnB stations (Figure 9).

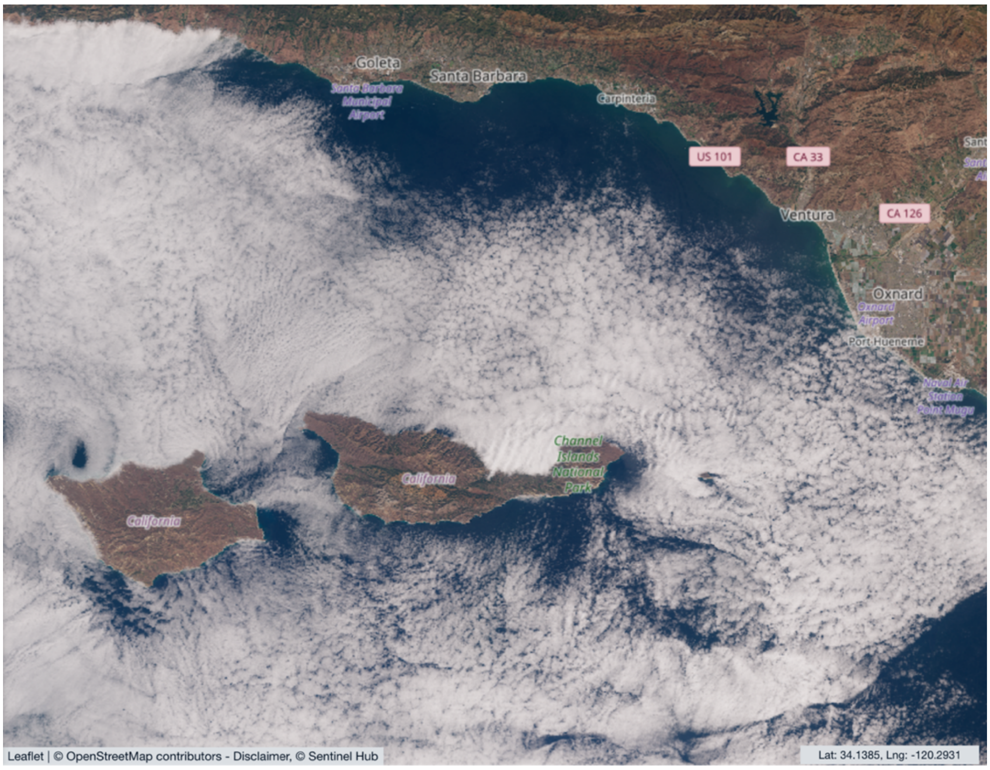

Figure 9: Screenshot taken from Copernicus Browser showing the presence of clouds on 24th September 2019 over SBC, thus limiting the extraction of chl-*a* values from some of the PnB stations.

Ideally, satellite-based imagery should have been used throughout the study to limit the variables between the data. However, due to time and technical constraints, satellite-based hyperspectral imagery was not possible for this study. AVIRIS data was selected instead because it has been operational since 1993 and because of its spatial resolution (20 m). However, AVIRIS imagery is also limited in its swath (12 km) compared to Sentinel-2 (290 km), and it is restricted to specific missions compared to Sentinel-2, which constantly captures global imagery.

Pre-processing of AVIRIS data introduced further complications. The large hyperspectral files required substantial computing power, leading to slow processing speeds and frequent software crashes. Licensing issues with the software added to the difficulties.

Moreover, applying the correct constants and band combinations was challenging, as the official algorithm documentation did not include specifications for the Sentinel-2 mission or AVIRIS instrument. This lack of direct documentation required reliance on estimates from other scientific journal articles, potentially introducing uncertainties into the chl-*a* concentration estimates. These limitations underscore the complexities and constraints encountered during the study, affecting data availability, processing efficiency, and the accuracy of the results.

## 4. Results

This section presents the findings of the hyperspectral and multispectral imagery analysis in estimating chl-*a* concentrations across the SBC (4.1), NB (4.2), and LoV (4.3) sites. The results are based on the analysis of remotely sensed data alongside *in-situ* measurements, with performance evaluated using metrics and statistical tests (4.4). Each site's results are detailed in the following subsections.

### 4.1. Santa Barbara Channel (SBC)

Chl-*a* concentrations at the SBC were measured through remotely sensed imagery and compared through *in-situ* measurements. The OC2 algorithm was applied to Sentinel-2 imagery on September 19th and 24th, 2019, and to AVIRIS imagery on September 25th, 2019. *In-situ* measurements from the PnB mission on September 18th provided a baseline for comparison.

On September 19th, the Sentinel-derived imagery displayed higher concentrations located in the northern stations (1-4) and lower values at the southern stations (5-7). A similar spatial pattern was observed on September 24th (Figure 10), although cloud cover limited data availability in the southern regions.

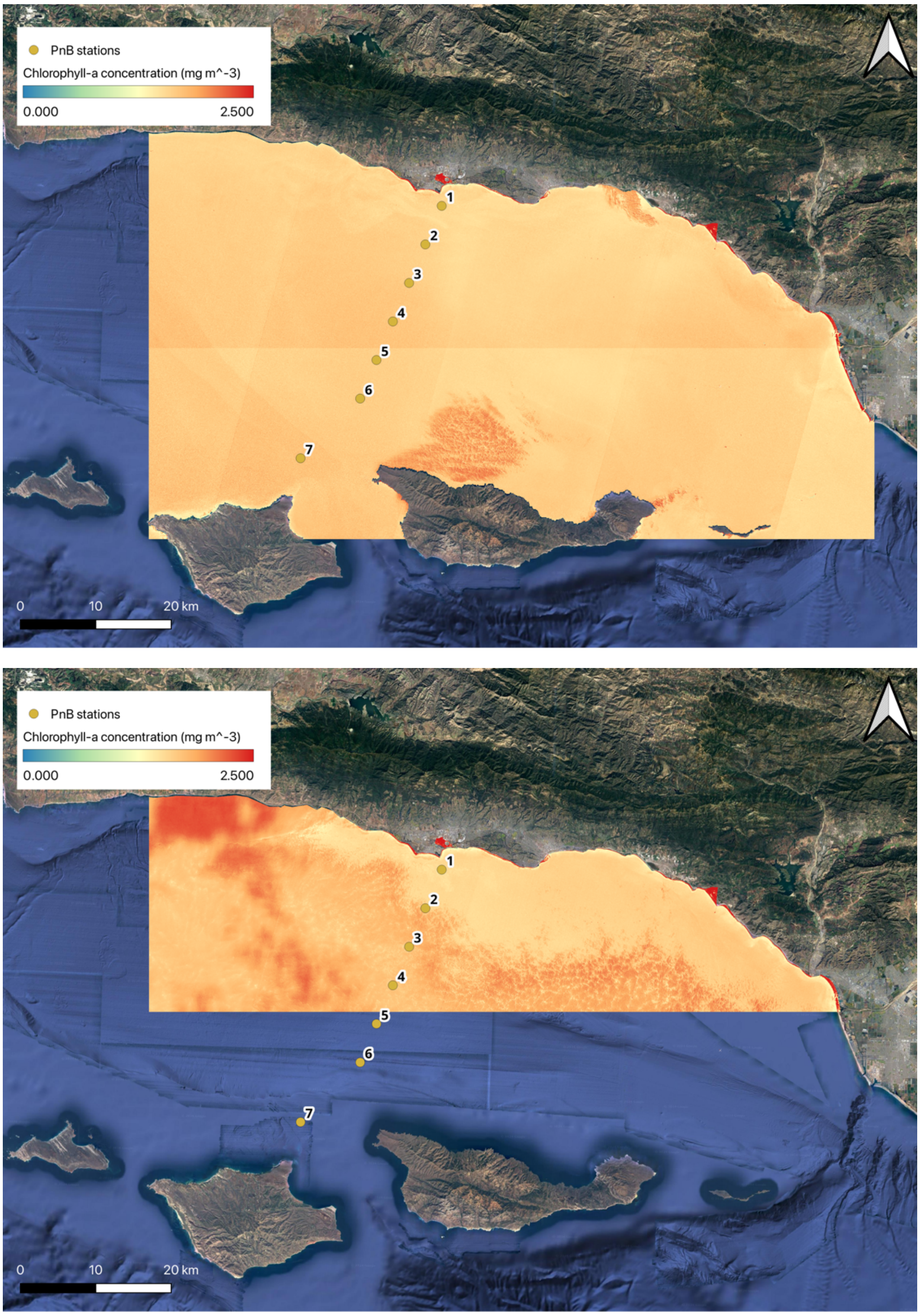

Figure 10: Output of the OC2 algorithm applied to Sentinel-2 imagery on 19th (top) and 24th (bottom) September 2019.

Similar to the Sentinel-2 outputs, the highest concentrations in thee AVIRIS imagery (Figure 11) were found in the northern coastal area (stations 1-4). However, the intensity was lower than the Sentinel-2 and *in-situ* data, particularly in the southern stations (5–7). Readings for station 6 were unavailable due to the narrow swath of the image capture.

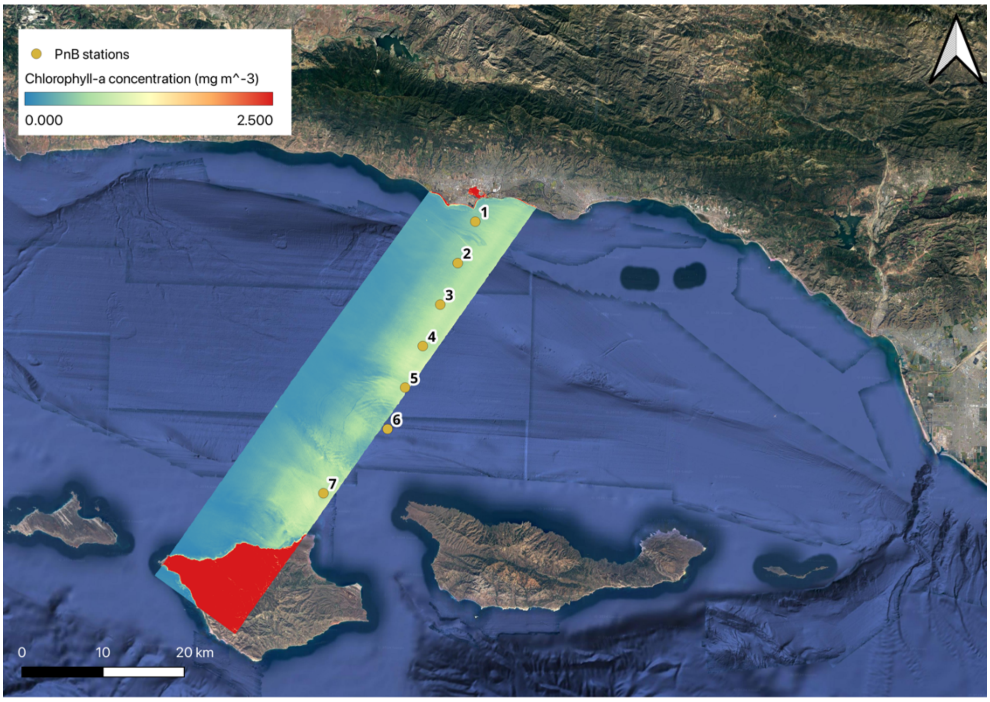

Figure 11: Outputs of the OC3 algorithm applied on AVIRIS imagery over SBC on 25th September 2019.

Table 2 summarises the chl-*a* concentrations at each station, highlighting a consistent decrease over time, especially for the AVIRIS data, which displayed the largest reduction in chl-*a* levels. Station 7 had the highest concentration in the PnB mission but experienced a significant decrease, possibly due to temporal dynamics.

Table 2: Chl-*a* concentration readings at the station from the PnB mission (18th September 2019) and the OCx algorithm applied on the Sentinel-2 imagery (19th & 24th September 2019) and AVIRIS imagery (25th September 2019).

| **Station** | **PnB chl-*a* (18/09)** | **Sentinel-2 chl-*a* (19/09)** | **Sentinel-2 chl-*a* (24/09)** | **AVIRIS chl-*a* (25/09)** |
|---|---|---|---|---|
| 1 | 2.06 | 1.63 | 1.42 | 0.499 |
| 2 | 1.37 | 1.58 | 1.93 | 0.542 |
| 3 | 1.42 | 1.68 | 1.59 | 0.712 |
| 4 | 0.954 | 1.78 | 1.72 | 0.849 |
| 5 | 0.992 | 1.67 | NaN | 0.831 |
| 6 | 1.05 | 1.67 | NaN | NaN |
| 7 | 2.16 | 1.75 | NaN | 0.670 |

## 4.2. Newport Bay (NB)

Chl-*a* concentrations for the NB AOI were derived from Sentinel-2 and AVIRIS imagery, with *in-situ* measurements from the USC_SEAPRISM station. Sentinel-2 imagery from June 27, 2023 (Figure 12), highlighted chl-*a* distribution across the region.

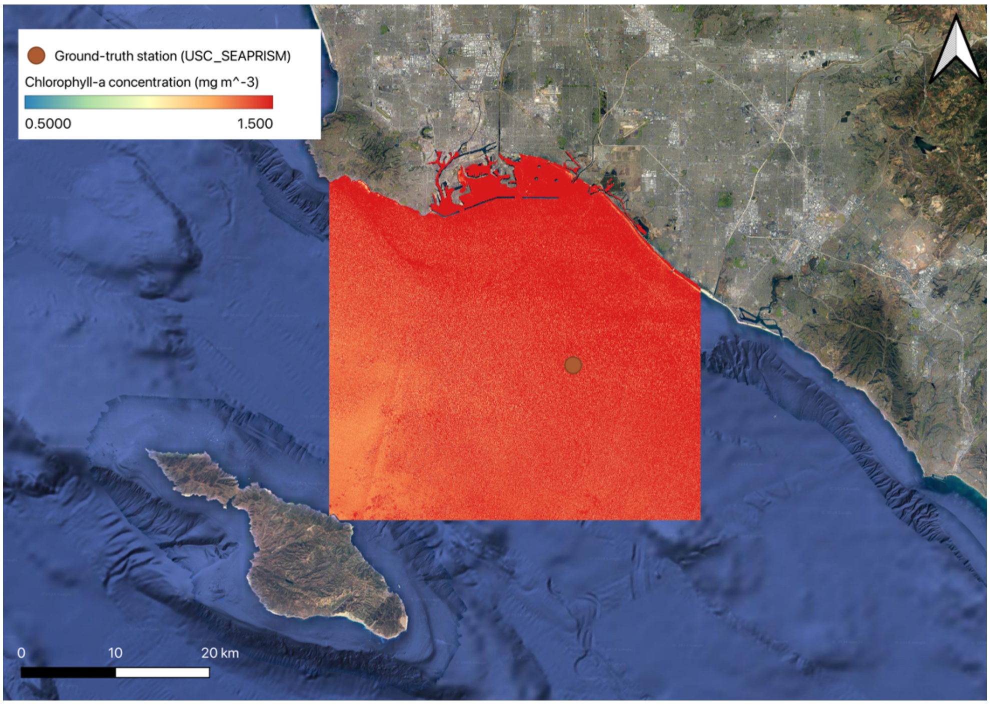

Figure 12: Outputs of the OC2 algorithm applied on Sentinel-2 imagery over NB on 27th June 2023.

The AVIRIS imagery, captured on June 28, 2023 (Figure 13), provided a more detailed spatial distribution of chl-*a* concentrations, a smaller range of concentrations.

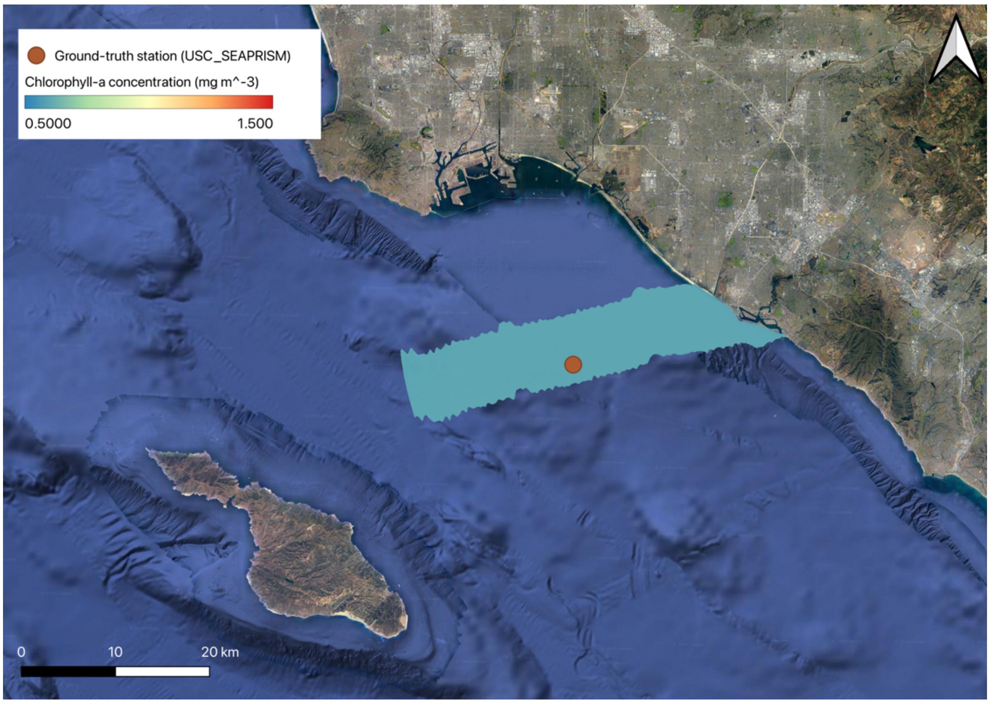

Figure 13: Outputs of the OC3 algorithm applied on AVIRIS imagery over NB on 28th June 2023.

The USC_SEAPRISM station recorded a chl-*a* concentration of 0.460 mg/m³, providing a reference for comparison with the remotely sensed data. Sentinel-2 recorded significantly higher chl-*a* concentrations (3.113 mg/m³) for the same station, while AVIRIS readings indicated a lower value (0.597 mg/m³) (Table 3).

Table 3: Chl-*a* concentration readings at the USC_SEAPRISM station (28 June 2023) and the OCx algorithm applied on the Sentinel-2 imagery (27 June 2023) and AVIRIS imagery (28 June 2023).

| **Station** | **USC_SEAPRISM chl-*a* (mg m$^{-3}$) (28/06)** | **Sentinel-2 chl-*a* (mg m$^{-3}$) (27/06)** | **AVIRIS chl-*a* (mg m$^{-3}$) (28/06)** |
|---|---|---|---|
| USC_SEAPRISM | 0.460 | 3.11 | 0.597 |

## 4.3. Lagoon of Venice (LoV)

Chl-*a* concentrations at the LoV site were obtained using Sentinel-2 and AVIRIS imagery from June 4, 2021, with ground-truth data from the AAOT station (Table 3). Sentinel-2 data shower chl-*a* concentration distribution (Figure 14), with higher levels near the northern coast and lower concentrations offshore.

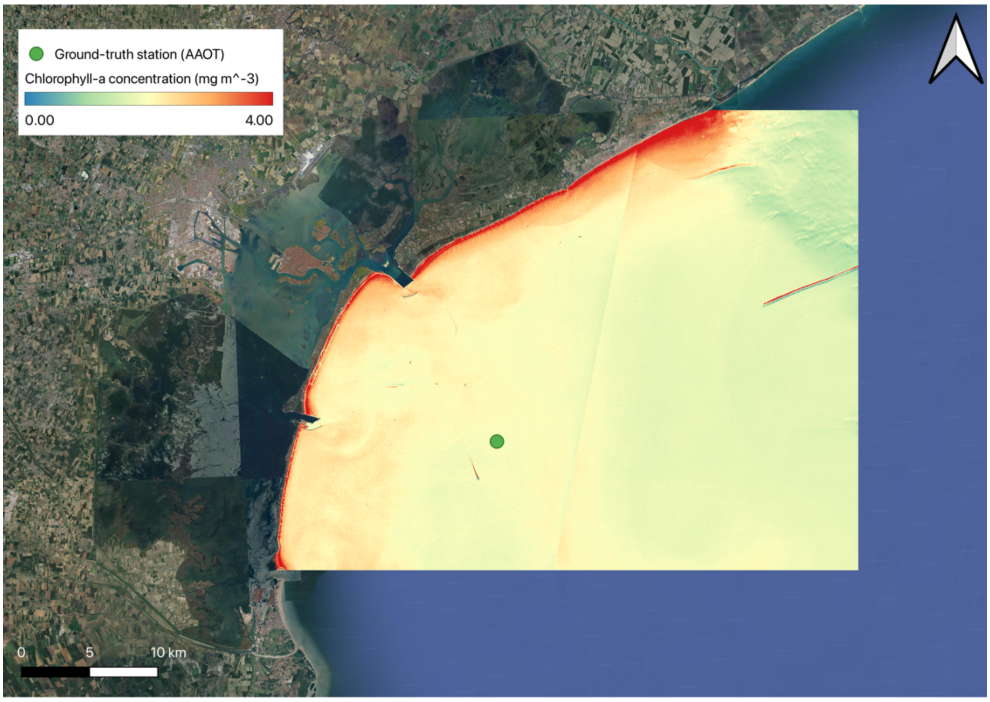

Figure 14: Outputs of the OC2 algorithm applied on Sentinel-2 imagery over LoV on 4th June 2021.

AVIRIS imagery collected on the same day (Figure 15) provided a more detailed spatial representation. This data showed higher concentrations along the northern coast, similar to the Sentinel-2 outputs.

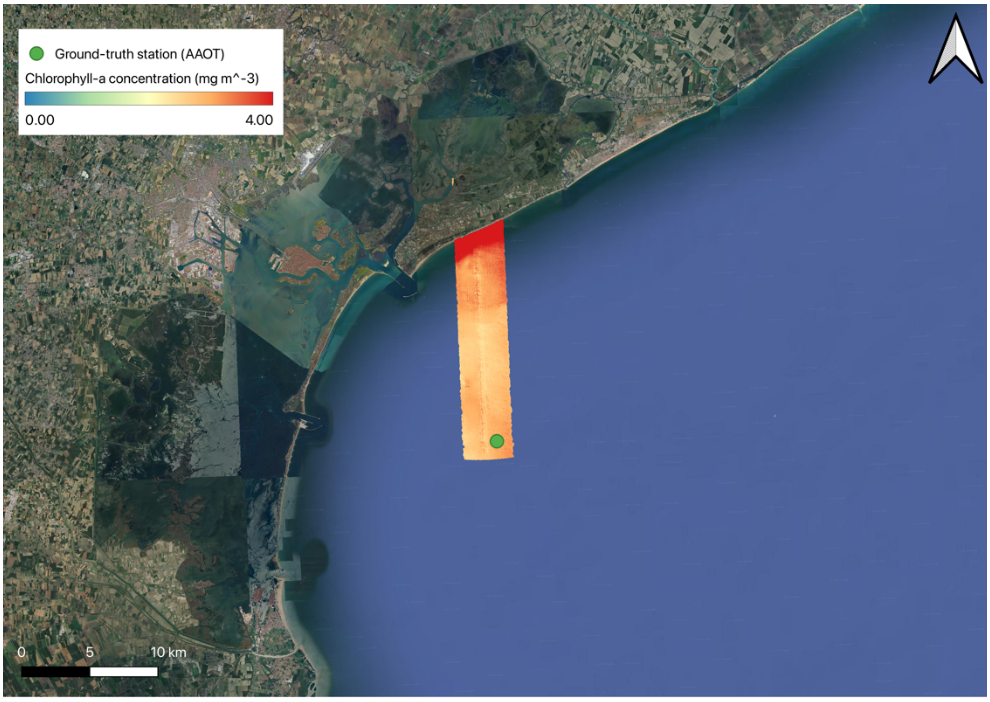

Figure 15: Outputs of the OC3 algorithm applied on AVIRIS imagery over LoV on 4th June 2021.

Ground-truth measurements from the AAOT station recorded a chl-*a* concentration of 3.34 mg/m³. The Sentinel-2 data at the same station showed a lower concentration of 2.09 mg/m³, while AVIRIS data recorded a slightly higher value of 3.10 mg/m³ (Table 4).

Table 4: Chl-*a* concentration readings from the AAOT station and the OCx algorithm applied on the Sentinel-2 and AVIRIS imagery (4th June 2021).

| **Station** | **AAOT chl-*a* (mg m$^{-3}$) (04/06)** | **Sentinel-2 chl-*a* (mg m$^{-3}$) (04/06)** | **AVIRIS chl-*a* (mg m$^{-3}$) (04/06)** |
|---|---|---|---|
| AAOT | 3.34 | 2.00 | 3.10 |

## 4.3. Comparison with ground-truth

The following section assesses the accuracy of chl-*a* concentration readings from Sentinel-2 and AVIRIS imagery at the three study sites: SBC (4.3.1.), NB (4.3.2.), and LoV (4.3.3.). The remotely sensed data are compared against ground-truth measurements, assessing the performance of each sensor and the algorithm applied.

### 4.3.1. Santa Barbara Channel (SBC)

The data from the SBC site was found to be normally distributed; paired t-tests were used to evaluate differences. Missing data caused by cloud coverage or image swath were handled using three methods: dropping rows with missing values, mean imputation, and linear interpolation. Sentinel-2 consistently demonstrated accuracy, with MAE values between 0.432 and 0.568 and RMSE values between 0.495 and 0.599. Paired t-tests indicated no significant differences from ground truth data ($p > 0.05$).

AVIRIS was initially less accurate with dropped rows (MAE = 0.801, RMSE = 0.954), but its performance  with mean imputation (MAE = 0.748, RMSE = 0.930) and linear interpolation (MAE = 0.739, RMSE = 0.927). Paired t-tests for AVIRIS indicated significant differences from the ground truth data ($p < 0.05$).

Overall, Sentinel-2 demonstrated reliable performance with low error values and no significant differences from the ground truth. Meanwhile, AVIRIS showed higher error values but improved significantly with data imputation methods (Table 5).

Table 5 : MAE, RMSE, and paired t-test results for Sentinel-2 and AVIRIS sensors at SBC, comparing three methods of handling missing data.

| **Site** | **Sensor** | **Date** | **Method** | **MAE** | **RMSE** | **Paired t-test (p value)** |
|---|---|---|---|---|---|---|
| SBC | Sentinel-2 | 19/09 | Dropped rows | 0.432 | 0.495 | 0.466 |
| | | 24/09 | | 0.534 | 0.579 | 0.544 |
| | AVIRIS | 25/09 | | 0.801 | 0.955 | 0.075 |
| | Sentinel-2 | 19/09 | Mean imputation | 0.492 | 0.535 | 0.246 |
| | | 24/09 | | 0.561 | 0.588 | 0.330 |
| | AVIRIS | 25/09 | | 0.748 | 0.930 | 0.016 |
| | Sentinel-2 | 19/09 | Linear interpolation | 0.492 | 0.535 | 0.246 |
| | | 24/09 | | 0.568 | 0.599 | 0.292 |
| | AVIRIS | 25/09 | | 0.739 | 0.927 | 0.018 |

### 4.3.2. Newport Bay (NB)

At NB, AVIRIS consistently demonstrated more precise results than Sentinel-2 in detecting chl-*a* concentrations. While the exact numerical metrics are omitted, AVIRIS achieved significantly better alignment with ground-truth values. This suggests that its finer spectral resolution offered an advantage in distinguishing chl-*a* from other pigments in coastal waters.

### 4.3.3. Lagoon of Venice (LoV)

For LoV, AVIRIS outperformed Sentinel-2 in chl-*a* concentration accuracy. Sentinel-2 demonstrated broader spectral band limitations, which hindered its ability to capture subtle variations in the water's properties. In contrast, AVIRIS exhibited enhanced detection of chl-*a* due to its hyperspectral capabilities, resulting in significantly better performance than Sentinel-2.

## 5. Discussion

This section highlights the insights from comparing hyperspectral and multispectral imagery in detecting chl-*a* concentrations across the three study sites (5.1). It discusses the strengths and limitations of each sensor's effectiveness and makes recommendations for future research (5.2).

### 5.1. Comparison between hyperspectral and multispectral imagery

The comparison between hyperspectral and multispectral imagery reveals differences in accuracy for detecting chl-*a* concentrations in coastal waters. While both Sentinel-2 and AVIRIS imagery provided valuable data, their performance varied depending on the site. The following sections analyse sensor performance at SBC (5.1.1), NB (5.1.2), and LoV (5.1.3).

### 5.1.1. Santa Barbara Channel (SBC)

Sentinel-2 demonstrated consistent performance at SBC, with relatively low error values regardless of the method used to handle missing data (e.g., MAE ~0.492 with imputation). Paired t-tests indicated no significant differences from ground-truth values ($p > 0.05$), affirming its reliability in predicting chl-*a* concentrations. These results suggest that Sentinel-2's multispectral data, while limited in spectral bands, can still deliver robust predictions in complex coastal environments like SBC.

In contrast, AVIRIS struggled to match Sentinel-2's performance at this site. It exhibited higher error metrics (e.g., MAE = 0.739 with interpolation) and significant differences from ground-truth values ($p = 0.018$ for interpolated data). Factors such as atmospheric conditions, variability in water composition, and the temporal mismatch between AVIRIS and ground-truth data (collected one week apart) may have influenced these results. Additionally, errors during the pre-processing stage (e.g., converting radiance to reflectance) might have further impacted AVIRIS performance.

Overall, Sentinel-2 outperformed AVIRIS at SBC, demonstrating that its coarser spectral resolution was adequate for capturing chl-*a* variability in this region.

### 5.1.2. Newport Bay (NB)

At NB, Sentinel-2 significantly overpredicted chl-*a* concentrations compared to ground-truth values. This discrepancy may be attributed to the challenges of using multispectral sensors in detecting a single CWQ parameter in highly dynamic environments. Additionally, using imagery from different days introduces temporal mismatches, which could have influenced the observed differences between sensor outputs and ground-truth data.

AVIRIS appeared to perform better at this site due to its finer spectral resolution, which may help differentiate chl-*a* from other pigments and CWQ parameters. However, this observation is based on a limited dataset, and the temporal mismatch between datasets further complicates direct comparisons. While these findings suggest a potential advantage for hyperspectral data in complex environments like NB, additional data collected on the same days and across multiple instances would be needed to substantiate this conclusion more confidently.

### 5.1.3. Lagoon of Venice (LoV)

At the LoV site, Sentinel-2 outputs appeared to underpredict chl-*a* concentrations, potentially indicating challenges in capturing small-scale variability in these measurements. This may be due to the sensor's broader and discrete spectral bands, which are less sensitive to subtle CWQ variations. However, it is important to note that this conclusion is based on a limited dataset, and further comparison with additional observations is required.

AVIRIS imagery appeared to perform better at this site, potentially due to its finer spectral resolution, which may enable more detailed detection of CWQ variations. The pre-processed AVIRIS data, already calibrated for reflectance, could have also influenced its outputs. While these findings suggest that hyperspectral sensors like AVIRIS may provide advantages in environments requiring detailed spectral data, additional studies and datasets would be necessary to confirm these results conclusively and to account for other influencing factors.

### 5.1.4. Overall findings

Across the three study sites, key trends emerged in the performance of hyperspectral and multispectral imagery for chl-*a* detection. Sentinel-2 demonstrated moderate accuracy but showed limitations in capturing small-scale variations in chl-*a* concentrations, particularly at the LoV and NB sites. These limitations may be attributed to its broader spectral bands, which reduce sensitivity to subtle CWQ variations. The complexity of coastal waters, combined with discrepancies in ground-truth data and coefficients used in the OCx algorithms, likely influenced these results. This was particularly evident at LoV and NB, where ground-truth comparisons were based on a single station, unlike the SBC site, which had a more robust sampling with seven stations.

AVIRIS, leveraging its hyperspectral capabilities, generally outperformed Sentinel-2 by providing better differentiation of chl-*a*. However, its performance was not consistent across all sites. For instance, at SBC, AVIRIS produced higher error values, which could be attributed to atmospheric disturbances, water complexity, and potential inaccuracies in band and coefficient selection during algorithm application. Additional challenges, such as pre-processing steps using the FLAASH tool to convert radiance to reflectance, may have introduced errors that impacted the results.

Despite these inconsistencies, AVIRIS demonstrated promise in detecting small-scale chl-*a* concentrations, particularly in areas requiring finer spectral resolution. These findings suggest that hyperspectral sensors hold significant potential for improving CWQ monitoring in complex environments. However, further refinement of processing algorithms, improved band selection, and expanded ground-truth comparison are necessary to strengthen these results and fully realise the advantages of hyperspectral imaging in such applications.

## 5.2. Recommendations for future studies

The outcomes of this study give rise to several recommendations for future research aiming to refine and improve the accuracy of chl-*a* concentration detection using remote sensing techniques.

Firstly, researchers should test different algorithms or adjust the OCx algorithms used in this study by experimenting with various coefficients and band combinations. Algorithms designed for optically complex waters may help improve accuracy and reduce discrepancies between multispectral and hyperspectral data and actual ground-truth readings.

Another recommendation is to temporally align *in-situ* data collection with remotely sensed imagery. In this study, data was collected on different days (due to availability), potentially leading to discrepancies due to the temporal variability of chl-*a* concentrations, which fluctuate rapidly in response to environmental changes. Aligning *in-situ* measurements with satellite or airborne acquisitions may mitigate this issue and improve accuracy.

As hyperspectral satellite imagery becomes more accessible, future studies should prioritise its use. While hyperspectral data from satellites was initially considered, its limited availability and spatial resolution at the time of this study restricted its use. However, the commercialisation and democratisation of hyperspectral imagery – driven by advancements in small-sat missions, cheaper components, and lighter sensors - are expected to provide accessible, higher-resolution hyperspectral data. As this imagery becomes more widely available, it could significantly enhance the ability to monitor chl-*a* concentrations with greater spatial coverage and spectral detail.

Lastly, future research should continue focusing on the importance of high spatial, spectral, and temporal resolutions in data collection. While hyperspectral data provides better spectral resolution, its less frequent revisit time, narrow swath, and lower spatial resolution can limit its ability to detect subtle chl-*a* level changes. In contrast, satellite-based multispectral imagery excels in frequent revisit times, wide swaths, and fine spatial resolutions while being easily accessible and, as in the case of Sentinel-2 data, often free of charge. A multi-sensor approach combining hyperspectral and multispectral imagery would address these issues, offering a better understanding of chl-*a* dynamics in coastal waters. This integration also supports the detection of HABs, where HSI's improved ability to differentiate between phytoplankton species and MSI's temporal and spatial resolutions could lead to more accurate HAB detection, mitigating negative impacts on public health and ecosystems.

## 6. Conclusion

This dissertation investigated the effectiveness of multispectral and hyperspectral remote sensing techniques for detecting and quantifying chl-*a* concentrations in three diverse coastal environments: the SBC, NB, and the LoV. By comparing Sentinel-2 and AVIRIS imagery with ground-truth data, the study aimed to assess the accuracy and utility of these technologies for CWQ monitoring, focusing on chl-*a* concentrations.

The results showed that Sentinel-2 provided moderate accuracy, particularly at the SBC site, where more robust ground-truth sampling allowed for better evaluation. However, its broader spectral bands limited its ability to capture fine-scale chl-*a* variability at NB and LoV. These limitations highlight the challenges MSI sensors face in dynamic and complex coastal environments.

In contrast, AVIRIS, with its finer spectral resolution, demonstrated stronger performance at NB and LoV, emphasising the advantages of HSI technology in detecting subtle variations in chl-*a* concentrations. However, AVIRIS's results at SBC were less accurate, potentially due to environmental factors, temporal mismatches with ground-truth data, or errors introduced during pre-processing or algorithm application.

Key insights emerged from this study. First, aligning *in-situ* data collection with remotely sensed imagery proved crucial, as temporal discrepancies likely contributed to inconsistencies, such as those observed at SBC. Second, refining algorithms and band selection is essential to enhance the accuracy and reliability of chl-*a* detection, particularly when using HSI data. These findings underscore the importance of continued advancements in sensor technology and processing techniques.

Overall, while both MSI and HSI sensors show promise for CWQ monitoring, HSI demonstrates a distinct advantage in environments where spectral complexity challenges MSI capabilities. With anticipated advancements in satellite-based HSI sensors, improved algorithms, and more robust data integration, the potential for large-scale CWQ monitoring and HAB mitigation is significant. These innovations will be critical for identifying nutrient-enriched waters and mitigating their adverse effects on public health and sensitive ecosystems.

## References

Abu Daya, M. (2004). *Coastal water quality monitoring with remote sensing in (East Kalimantan) Makassar Strait, Indonesia* (MSc. in Geo-Information Science and Earth Observation, Coastal Zones Studies). International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands. Retrieved from https://webapps.itc.utwente.nl/librarywww/papers_2004/msc/ereg/abu_ daya.pdf

Acri, F., Braga, F., & Aubry, F. B. (2020). Long-term dynamics in nutrients, chlorophyll a and water quality parameters in the Lagoon of Venice. *Scientia Marina*, *84*(3), 215. https://doi.org/10.3989/scimar.05022.30a

AERONET, Jones, B., & Ragan, M. (2024). *Ocean Color (OC) Data Display Interface: USC_SEAPRISM* [Dataset]. Retrieved from https://aeronet.gsfc.nasa.gov/cgi-bin/data_display_seaprism_v4013?site=USC_SEAPRISM&nachal=2&level=2&place_code=10

AERONET, Zibordi, G., & Bulgarelli, B. (2024). *Ocean Color (OC) Data Display Interface: AAOT* [Dataset]. Retrieved from https://aeronet.gsfc.nasa.gov/cgi-bin/data_display_seaprism_v4013?site=AAOT&nachal=2&level=2&place_code=10

Anderson, D. M. (2009). Approaches to monitoring, control and management of harmful algal blooms (HABs). *Ocean & Coastal Management*, *52*(7), 342–347. https://doi.org/10.1016/j.ocecoaman.2009.04.006

Anderson, E. L., Duranleau, L., Garrison, G. E., & Valachovic, R. W. (2015). U.S. Dental School Applicants and Enrollees, 2013 Entering Class. *Journal of Dental Education*, *79*(5), 592–605. https://doi.org/10.1002/j.0022-0337.2015.79.5.tb05919.x

Aubry, F. B., Acri, F., Finotto, S., & Pugnetti, A. (2021). Phytoplankton dynamics and water quality in the Venice Lagoon. *Water*, *13*(19), 2780. https://doi.org/10.3390/w13192780

Beers, J., Trent, J., Reid, F., & Shanks, A. (1986). Macroaggregates and their phytoplanktonic components in the Southern California Bight. *Journal of Plankton Research*, *8*(3), 475–487. https://doi.org/10.1093/plankt/8.3.475

Bian, C., Jiang, W., & Song, D. (2010). Terrigenous transportation to the Okinawa Trough and the influence of typhoons on suspended sediment concentration. *Continental Shelf Research*, *30*(10–11), 1189–1199. https://doi.org/10.1016/j.csr.2010.03.008

Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., & Ruescas, A. (2016). Evolution of the C2RCC neural Network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. In L. Ouwehand (Ed.), *ESA-SP* (Vol. 740, p. 54). Prague, Czechia. Retrieved from https://articles.adsabs.harvard.edu/pdf/2016ESASP.740E..54B

California State boundary. (2018). [Dataset]. *ArcGIS Hub*. Retrieved from https://hub.arcgis.com/datasets/febda4ff242f4f699dc40183833ea141/

Çevirgen, S., Elwany, H., Pesce, M., & Zirino, A. (2020). Managing nutrient pollution in Venice Lagoon (Italy): a practical tool for assessment of water quality. *Sustainable Water Resources Management*, *6*(3). https://doi.org/10.1007/s40899-020-00390-y

Copernicus. (n.d.). SentiWiki. Retrieved from https://sentiwiki.copernicus.eu/web/sentiwiki

Dekker, A. G., Pinnel, N., Gege, P., Briottet, X., Court, A., Peters, S., . . . Pflug, B. (2018, January 23). Feasibility Study for an Aquatic Ecosystem Earth Observing System Version 1.2. Retrieved from https://hal.science/hal-02172188/

Dierssen, H. M., Ackleson, S. G., Joyce, K. E., Hestir, E. L., Castagna, A., Lavender, S., & McManus, M. A. (2021). Living up to the hype of hyperspectral aquatic remote sensing: Science, resources and outlook. *Frontiers in Environmental Science*, *9*. https://doi.org/10.3389/fenvs.2021.649528

Dorman, C. E., & Winant, C. D. (2000). The Structure and Variability of the Marine Atmosphere around the Santa Barbara Channel. *Monthly Weather Review*, *128*(2), 261. https://doi.org/10.1175/1520-0493(2000)128

Duan, W., He, B., Takara, K., Luo, P., Nover, D., Sahu, N., & Yamashiki, Y. (2013). Spatiotemporal evaluation of water quality incidents in Japan between 1996 and 2007. *Chemosphere*, *93*(6), 946–953. https://doi.org/10.1016/j.chemosphere.2013.05.060

Duan, W., Takara, K., He, B., Luo, P., Nover, D., & Yamashiki, Y. (2013). Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari River, Japan, 1985 to 2010. *Science of the Total Environment*, *461–462*, 499–508. https://doi.org/10.1016/j.scitotenv.2013.05.022

ENVI Tutorials. (n.d.). Preprocessing AVIRIS data Tutorial. Retrieved from https://www.nv5geospatialsoftware.com/docs/PreprocessAVIRIS.html

Environmental Protection Agency [EPA]. (2024, May 10). Indicators: chlorophyll-a. Retrieved from https://www.epa.gov/national-aquatic-resource-surveys/indicators-chlorophyll

Eppley, R. W., Renger, E. H., & Harrison, W. G. (1979). Nitrate and phytoplankton production in southern California coastal waters1. *Limnology and Oceanography/˜the œL & O on Cd-rom*, *24*(3), 483–494. https://doi.org/10.4319/lo.1979.24.3.0483

Erwin, S. (2023, March 22). NRO signs agreements with commercial providers of hyperspectral imagery. Retrieved from https://spacenews.com/nro-signs-agreements-with-commercial-providers-of-hyperspectral-imagery/

European Commission. (2000). *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy (OJ L 327 22.12.2000 p. 1)* (pp. 879–969). Retrieved from https://eur-lex.europa.eu/eli/dir/2000/60/2014-11-20

European Environment Agency [EEA]. (1991). Coastal waters. Retrieved from https://www.eea.europa.eu/help/glossary/eea-glossary/coastal-waters

European Environment Agency [EEA]. (2024, June 13). Bathing water quality. Retrieved from https://www.eea.europa.eu/en/topics/in-depth/bathing-water

European Space Agency. (n.d.). SNAP. Retrieved from https://earth.esa.int/eogateway/tools/snap

European Space Agency [ESA]. (2020). Types of orbits. Retrieved from https://www.esa.int/Enabling_Support/Space_Transportation/Types_of_orbits

European Space Agency [ESA]. (n.d.). Level-1C Processing - Sentinel-2 MSI Technical Guide. Retrieved from https://sentiwiki.copernicus.eu/web/s2-processing

Facca, C., Bilaničovà, D., Pojana, G., Sfriso, A., & Marcomini, A. (2014). Harmful algae records in Venice Lagoon and in Po River Delta (Northern Adriatic Sea, Italy). *The Scientific World JOURNAL*, *2014*, 1–11. https://doi.org/10.1155/2014/806032

Fernández-Tejedor, M., Velasco, J. E., & Angelats, E. (2022). Accurate estimation of chlorophyll-A concentration in the coastal areas of the Ebro Delta (NW Mediterranean) using Sentinel-2 and its application in the selection of areas for mussel aquaculture. *Remote Sensing*, *14*(20), 5235. https://doi.org/10.3390/rs14205235

Gholizadeh, M. H., Melesse, A. M., & Reddi, L. (2016). A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors*, *16*(8), 1298. https://doi.org/10.3390/s16081298

Gorgoglione, A., Gioia, A., Iacobellis, V., Piccinni, A. F., & Ranieri, E. (2016). A rationale for pollutograph evaluation in ungauged areas, using daily rainfall patterns: case studies of the Apulian region in southern Italy. *Applied and Environmental Soil Science*, *2016*, 1–16. https://doi.org/10.1155/2016/9327614

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., . . . Williams, O. (1998). Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, *65*(3), 227–248. https://doi.org/10.1016/s0034-4257(98)00064-9

Gregor, J., & Maršálek, B. (2004). Freshwater phytoplankton quantification by chlorophyll a: a comparative study of in vitro, in vivo and in situ methods. *Water Research*, *38*(3), 517–522. https://doi.org/10.1016/j.watres.2003.10.033

Gruber, N., & Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature*, *451*(7176), 293–296. https://doi.org/10.1038/nature06592

Hendershott, N., & Winant, N. (1996). Surface circulation in the Santa Barbara Channel. *Oceanography*, *9*(2), 114–121. https://doi.org/10.5670/oceanog.1996.14

Hickey, B. M. (1979). The California current system—hypotheses and facts. *Progress in Oceanography/Progress in Oceanography*, *8*(4), 191–279. https://doi.org/10.1016/0079-6611(79)90002-8

Holt, B., Trinh, R., & Gierach, M. M. (2017). Stormwater runoff plumes in the Southern California Bight: A comparison study with SAR and MODIS imagery. *Marine Pollution Bulletin*, *118*(1–2), 141–154. https://doi.org/10.1016/j.marpolbul.2017.02.040

Ignaciov. (2022, November 6). Hyperspectral and multispectral imaging. Retrieved from https://www.pyroistech.com/hyperspectral-and-multispectral-imaging/

Jones, B. H., Atkinson, L. P., Blasco, D., Brink, K. H., & Smith, S. L. (1988). The asymmetric distribution of chlorophyll associated with a coastal upwelling center. *Continental Shelf Research*, *8*(10), 1155–1170. https://doi.org/10.1016/0278-4343(88)90017-9

Kamer, K., Boyle, K. A., & Fong, P. (2001). Macroalgal bloom dynamics in a highly eutrophic Southern California estuary. *Estuaries*, *24*(4), 623. https://doi.org/10.2307/1353262

Kennison, R., Kamer, K., & Fong, P. (2003, December). Nutrient dynamics and macroalgal blooms: A comparison of five Southern California estuaries. *Southern California*

*Coastal Water Research Project*. Southern California Coastal Water Research Project. Retrieved from https://www.researchgate.net/publication/228780670_Nutrient_dynamics_and_macroalgal_blooms_a_comparison_of_five_southern_California_estuaries

Kolluru, S., Gedam, S. S., Chander, S., & Sahay, A. (2023). Development of chlorophyll-a concentration estimation algorithm for turbid productive inland waters in India. *Geocarto International*, *38*(1). https://doi.org/10.1080/10106049.2023.2171143

Levin, L. A., Liu, K. K., Emeis, K., Breitburg, D., Cloern, J. E., Deutsch, C., . . . Wishner, K. F. (2015). Comparative biogeochemistry–ecosystem–human interactions on dynamic continental margins. *Journal of Marine Systems*, *141*, 3–17. https://doi.org/10.1016/j.jmarsys.2014.04.016

Lock, M., Saintilan, N., Van Duren, I., & Skidmore, A. (2023). Monitoring Coastal Water Body Health with Sentinel-2 MSI Imagery. *Remote Sensing*, *15*(7), 1734. https://doi.org/10.3390/rs15071734

Louis, V., Tanti, C. M., & Rolé, A. (2005). Coastal area management in Malta. *Priority Actions Programme for UNEP/MAP*. Retrieved from https://discomap.eea.europa.eu/map/Data/Milieu/OURCOAST_032_MT/OURCOAST_032_MT_Doc1_MTCoastManagementUNEP.pdf

Lunetta, R. S., Knight, J. F., Paerl, H. W., Streicher, J. J., Peierls, B. L., Gallo, T., . . . Buzzelli, C. P. (2009). Measurement of water colour using AVIRIS imagery to assess the potential for an operational monitoring capability in the Pamlico Sound Estuary, USA. *International Journal of Remote Sensing*, *30*(13), 3291–3314. https://doi.org/10.1080/01431160802552801

Lynn, R. J., & Simpson, J. J. (1987). The California Current system: The seasonal variability of its physical characteristics. *Journal of Geophysical Research*, *92*(C12), 12947–12966. https://doi.org/10.1029/jc092ic12p12947

Mertes, L. A., & Warrick, J. A. (2001). Measuring flood output from 110 coastal watersheds in California with field measurements and SeaWiFS. *Geology*, *29*(7), 659. https://doi.org/10.1130/0091-7613(2001)029

Moita, M., Sobrinho-Gonçalves, L., Oliveira, P., Palma, S., & Falcão, M. (2006). A bloom ofDinophysis acutain a thin layer off North-West Portugal. *African Journal of Marine Science*, *28*(2), 265–269. https://doi.org/10.2989/18142320609504160

Moses, W. J., Gitelson, A. A., Berdnikov, S., Bowles, J. H., Povazhnyi, V., Saprygin, V., . . . Patterson, K. W. (2014). HICO-Based NIR–Red models for estimating chlorophyll-$A$ concentration in productive coastal waters. *IEEE Geoscience and Remote Sensing Letters*, *11*(6), 1111–1115. https://doi.org/10.1109/lgrs.2013.2287458

Moses, W. J., Gitelson, A. A., Perk, R. L., Gurlin, D., Rundquist, D. C., Leavitt, B. C., . . . Brakhage, P. (2012). Estimation of chlorophyll-a concentration in turbid productive waters using airborne hyperspectral data. *Water Research*, *46*(4), 993–1004. https://doi.org/10.1016/j.watres.2011.11.068

Muller-Karger, F. E., Hestir, E., Ade, C., Turpie, K., Roberts, D. A., Siegel, D., . . . Jetz, W. (2018). Satellite sensor requirements for monitoring essential biodiversity variables of coastal ecosystems. *Ecological Applications*, *28*(3), 749–760. https://doi.org/10.1002/eap.1682

Municipal Boundaries of Italy 2019. (2019). [Dataset]. *ArcGIS Hub*. Retrieved from https://hub.arcgis.com/datasets/inspire-esri::municipal-boundaries-of-italy-2019/

NASA Jet Propulsion Laboratory. (n.d.-a). AVIRIS-Classic Data Portal. Retrieved from https://aviris.jpl.nasa.gov/dataportal/

NASA Jet Propulsion Laboratory. (n.d.-b). Index of /pub/cal_files/. Retrieved from https://popo.jpl.nasa.gov/pub/cal_files/

Nazeer, M., Bilal, M., Alsahli, M., Shahzad, M., & Waqas, A. (2017). Evaluation of empirical and machine learning algorithms for estimation of coastal water quality parameters. *ISPRS International Journal of Geo-information*, *6*(11), 360. https://doi.org/10.3390/ijgi6110360

Nezlin, N. P., Kamer, K., Hyde, J., & Stein, E. D. (2009). Dissolved oxygen dynamics in a eutrophic estuary, Upper Newport Bay, California. *Estuarine Coastal and Shelf Science*, *82*(1), 139–151. https://doi.org/10.1016/j.ecss.2009.01.004

Nezlin, N. P., Kamer, K., & Stein, E. D. (2007). Application of color infrared aerial photography to assess macroalgal distribution in an eutrophic estuary, Upper Newport Bay, California. *Estuaries and Coasts*, *30*(5), 855–868. https://doi.org/10.1007/bf02841339

Niroumand-Jadidi, M., Bovolo, F., Bruzzone, L., & Gege, P. (2021). Inter-Comparison of Methods for Chlorophyll-a retrieval: Sentinel-2 Time-Series Analysis in Italian lakes. *Remote Sensing*, *13*(12), 2381. https://doi.org/10.3390/rs13122381

Olmanson, L. G., Brezonik, P. L., & Bauer, M. E. (2013). Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. *Remote Sensing of Environment*, *130*, 254–265. https://doi.org/10.1016/j.rse.2012.11.023

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

Otero, M., & Siegel, D. (2004). Spatial and temporal characteristics of sediment plumes and phytoplankton blooms in the Santa Barbara Channel. *Deep-sea Research. Part 2.*

*Topical Studies in Oceanography/Deep Sea Research. Part II, Topical Studies in Oceanography*, *51*(10–11), 1129–1149. https://doi.org/10.1016/j.dsr2.2004.04.004

Paerl, H. W., Valdes, L. M., Pinckney, J. L., Piehler, Dyble, J., & Moisander, P. H. (2003). Phytoplankton photopigments as indicators of estuarine and coastal eutrophication. *BioScience*, *53*(10), 953–964. https://doi.org/10.1641/0006-3568

Palma, S., Mouriño, H., Silva, A., Barão, M. I., & Moita, M. T. (2010). Can Pseudo-nitzschia blooms be modeled by coastal upwelling in Lisbon Bay? *Harmful Algae*, *9*(3), 294–303. https://doi.org/10.1016/j.hal.2009.11.006

Pinckney, J. L., Paerl, H. W., Harrington, M. B., & Howe, K. E. (1998). Annual cycles of phytoplankton community-structure and bloom dynamics in the Neuse River Estuary, North Carolina. *Marine Biology*, *131*(2), 371–381. https://doi.org/10.1007/s002270050330

Poggioli, S. (2008, January 7). MOSE project aims to part Venice floods. *NPR*. Retrieved from https://www.npr.org

Ritchie, J. C., & Schiebe, F. R. (1998). Water quality. In E. T. Engman & G. A. Schultz (Eds.), *Remote Sensing in Hydrology and Water Management* (pp. 287–303). Heidelberg, Germany: Springer. https://doi.org/10.1007/978-3-642-59583-7_13

Ritchie, J. C., Zimba, P. V., & Everitt, J. H. (2003). Remote sensing techniques to assess water quality. *Photogrammetric Engineering and Remote Sensing*, *69*(6), 695–704. https://doi.org/10.14358/pers.69.6.695

SCB Marine Biodiversity Observation Network, Catlett, D., Siegel, D. A., Guillocheau, N., & Kui, L. (2024). *Plumes and Blooms: phytoplankton pigment concentration* [Dataset]. https://doi.org/10.6073/pasta/bf1826d73714586b0fbc2875a5420898

Palacios, S. L., Kudela, R. M., Guild, L. S., Negrey, K. H., Torres-Perez, J., & Broughton, J. (2015). Remote sensing of phytoplankton functional types in the coastal ocean from the HyspIRI Preparatory Flight Campaign. *Remote Sensing of Environment*, *167*, 269–280. https://doi.org/10.1016/j.rse.2015.05.014

Silva, A., Palma, S., Oliveira, P., & Moita, M. (2009). Composition and interannual variability of phytoplankton in a coastal upwelling region (Lisbon Bay, Portugal). *Journal of Sea Research*, *62*(4), 238–249. https://doi.org/10.1016/j.seares.2009.05.001

Silva, A., Pinto, L., Rodrigues, S., De Pablo, H., Santos, M., Moita, T., & Mateus, M. (2016). A HAB warning system for shellfish harvesting in Portugal. *Harmful Algae*, *53*, 33–39. https://doi.org/10.1016/j.hal.2015.11.017

Smayda, T. J. (2002). Turbulence, watermass stratification and harmful algal blooms: an alternative view and frontal zones as “pelagic seed banks.” *Harmful Algae*, *1*(1), 95–112. https://doi.org/10.1016/s1568-9883(02)00010-0

Stumpf, R. P., Tomlinson, M. C., Calkins, J. A., Kirkpatrick, B., Fisher, K., Nierenberg, K., . . . Wynne, T. T. (2010). Skill assessment for an operational algal bloom forecast system. *Journal of Marine Systems*, *76*(1–2), 151–161. https://doi.org/10.1016/j.jmarsys.2008.05.016

Supported missions: CZCS. (n.d.). Retrieved from https://oceancolor.gsfc.nasa.gov/about/missions/czcs/

Trainer, V., Pitcher, G., Reguera, B., & Smayda, T. (2010). The distribution and impacts of harmful algal bloom species in eastern boundary upwelling systems. *Progress in Oceanography/Progress in Oceanography*, *85*(1–2), 33–52. https://doi.org/10.1016/j.pocean.2010.02.003

Vandermeulen, R. A., Mannino, A., Craig, S. E., & Werdell, P. J. (2020). 150 shades of green: Using the full spectrum of remote sensing reflectance to elucidate color shifts in the ocean. *Remote Sensing of Environment*, *247*, 111900. https://doi.org/10.1016/j.rse.2020.111900

Vandermeulen, R. A., Mannino, A., Neeley, A., Werdell, J., & Arnone, R. (2017). Determining the optimal spectral sampling frequency and uncertainty thresholds for hyperspectral remote sensing of ocean color. *Optics Express*, *25*(16), A785. https://doi.org/10.1364/oe.25.00a785

Venrick, E. L. (1998). The phytoplankton of the Santa Barbara Basin: patterns of chlorophyll and species structure and their relationships with those of surrounding stations. *CalCOFl Rep.* Retrieved from http://www.calcofi.org/publications/calcofireports/v39/Vol_39_Venrick.pdf

Werdell, J., O'Reilly, J., Hu, C., Feng, L., Lee, Z., Franz, B., . . . Proctor, C. (2023). Chlorophyll a [Dataset]. (G. Wang, Ed.), *NASA Earthdata: Algorithm Publication Tool Beta*. https://doi.org/10.5067/JCQB8QALDOYD

Wolanin, A., Soppa, M. A., & Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sensing*, *8*(10), 871. https://doi.org/10.3390/rs8100871

Wolff, J., Grashorn, S., Lettmann, K., & Stanev, E. V. (2014). East-Frisian Wadden Sea hydrodynamics and wave effects in an unstructured-grid model. *ResearchGate*. https://doi.org/10.13140/2.1.5055.3126

World Health Organization [WHO]. (2006). 9. Aesthetic issues. In *Guidelines for Safe Recreational Water Environments* (pp. 159–167). Retrieved from https://www.who.int/water_sanitation_health/bathing/srwe1- chap9.pdf

Appendix

1. Coordinates of the PnB stations

| Station | Longitude | Latitude |
|---|---|---|
| 1 | 34,39017 | -119,84067 |
| 2 | 34,3435 | -119,86267 |
| 3 | 34,29683 | -119,8845 |
| 4 | 34,25017 | -119,90633 |
| 5 | 34,2035 | -119,92833 |
| 6 | 34,15683 | -119,95017 |
| 7 | 34,08333 | -120,03333 |

2. Python script to apply OC2 algorithm to SENTINEL-2 imagery

```
import rasterio
import numpy as np

# Coefficients for OC3_VIIRS_SNPP
coeffs =[0.3308, -2.66837, 1.59899, 0.55248, -1.4876]
# Function to calculate chlorophyll-a concentration using OC3_VIIRS_SNPP
def calculate_chlor_a(Rrs_490, Rrs_560):
    ratio = np.log10(np.maximum(Rrs_490) / Rrs_560)
    chlor_a_log = coeffs[0]
    for i in range(1, 5):
        chlor_a_log += coeffs[i] * (ratio ** i)
    chlor_a = 10 ** chlor_a_log
    return chlor_a

# Path to output file
output_file = '/outputs/L2A_20230627_ocx.tif'

# Read the input raster
with rasterio.open(' /L2A_20230627_resampled_msk_B2B3B4.tif') as src:
    # Read the relevant bands (assuming band 1 is Rrs_490, band 2 is Rrs_560)
    Rrs_490 = src.read(1)
    Rrs_560 = src.read(2)

    # Calculate chlorophyll-a concentration for each pixel
    chlor_a = calculate_chlor_a(Rrs_490, Rrs_560)

    # Define the metadata for the output file
    profile = src.profile
    profile.update(dtype=rasterio.float32, count=1)

    # Write the output raster
```

```
    with rasterio.open(output_file, 'w', **profile) as dst:
        dst.write(chlor_a.astype(rasterio.float32), 1)

print(f"Chlorophyll-a concentration raster saved to {output_file}")
```

3. Python script to apply OC3 algorithm to AVIRIS imagery

```
import rasterio
import numpy as np
import pandas as pd  # Make sure to import pandas

# Coefficients for OC3_VIIRS_SNPP
coeffs =[0.2521,-2.2146,1.5193,-0.7702,-0.4921]

#R: 667.5nm, G: 540nm, B: 470nm
# Function to calculate chlorophyll-a concentration using OC3_VIIRS_SNPP
def calculate_chlor_a(Rrs_443, Rrs_486, Rrs_551):
    ratio = np.log10(np.maximum(Rrs_443, Rrs_486) / Rrs_551)
    chlor_a_log = coeffs[0]
    for i in range(1, 5):
        chlor_a_log += coeffs[i] * (ratio ** i)
    chlor_a = 10 ** chlor_a_log
    return chlor_a

# Define the bands DataFrame
bands = pd.DataFrame({
    "Band center (nm)": [443, 486, 551],
    "Band number": [14,23,36]
})

# Path output file
output_file = ' /outputs/ang20230628_ocx_.tif'

# Read the input raster
with rasterio.open("/ang20230628.tif") as src:
    # Function to get the band number closest to the given wavelength
    get_band_number = lambda w: bands.iloc[(bands["Band center (nm)"] -
w).abs().argsort()[0]]["Band number"]
```

```
    # Get band numbers for the specified wavelengths and ensure they are integers
    b_Rrs_443 = int(get_band_number(443))
    b_Rrs_486 = int(get_band_number(486))
    b_Rrs_551 = int(get_band_number(551))

    print(b_Rrs_443, b_Rrs_486, b_Rrs_551)

    # Read the relevant bands from the raster file
    Rrs_443 = src.read(b_Rrs_443).astype(float)  # Convert to float
    Rrs_486 = src.read(b_Rrs_486).astype(float)  # Convert to float
    Rrs_551 = src.read(b_Rrs_551).astype(float)  # Convert to float

    # Replace invalid data with NaN
    Rrs_443[Rrs_443 == -9999.] = np.nan
    Rrs_486[Rrs_486 == -9999.] = np.nan
    Rrs_551[Rrs_551 == -9999.] = np.nan

    # Calculate chlorophyll-a concentration for each pixel
    chlor_a = calculate_chlor_a(Rrs_443, Rrs_486, Rrs_551)

    # Define the metadata for the output file
    profile = src.profile
    profile.update(dtype=rasterio.float32, count=1, driver='GTiff')

    # Write the output raster
    with rasterio.open(output_file, 'w', **profile) as dst:
        dst.write(chlor_a.astype(rasterio.float32), 1)

print(f"Chlorophyll-a concentration raster saved to {output_file}")
```

### 4. Python script to perform statistical tests of results

```
import numpy as np
import pandas as pd
from scipy.stats import shapiro
from scipy import stats
from sklearn.metrics import mean_absolute_error, mean_squared_error

# Data/observations for LoV

LoV_AAOT = [3.342]  # LoV has only one ground-truth value
LoV_Sentinel = [2.085]       # Predicted value from Sentinel imagery for LoV
LoV_AVIRIS = [3.103] # Predicted value from AVIRIS imagery for LoV

# Calculating MAE & RMSE for LoV

mae_sentinel_lov = mean_absolute_error([LoV_AAOT], [LoV_Sentinel])
mae_aviris_lov = mean_absolute_error([LoV_AAOT], [LoV_AVIRIS])

print(f"Mean Absolute Error for LoV (Sentinel): {mae_sentinel_lov:.4f}")
print(f"Mean Absolute Error for LoV (AVIRIS): {mae_aviris_lov:.4f}")

# Data/observations for LoV

Newport_USC = [0.4597]  # Newport has only one ground-truth value
Newport_Sentinel = [3.113]       # Predicted value from Sentinel imagery for Newport
Newport_AVIRIS = [0.5971] # Predicted value from AVIRIS imagery for Newport

# Calculating MAE & RMSE for NB

mae_sentinel_newport = mean_absolute_error([Newport_USC], [Newport_Sentinel])
mae_aviris_newport = mean_absolute_error([Newport_USC], [Newport_AVIRIS])
```

```
print(f"Mean Absolute Error for Newport (Sentinel): {mae_sentinel_newport:.4f}")
print(f"Mean Absolute Error for Newport (AVIRIS): {mae_aviris_newport:.4f}")

#SBC site: dealing with missing values/NaNs

# Create a DataFrame
data = {
   'Station': [1, 2, 3, 4, 5, 6, 7],
   'PnB_chl_a': [2.062, 1.365, 1.424, 0.954, 0.992, 1.048, 2.164],
   'Sentinel-2_chl_a_19_09': [1.628, 1.578, 1.677, 1.781, 1.674, 1.666, 1.745],
   'Sentinel-2_chl_a_24_09': [1.418, 1.930, 1.590, 1.715, np.nan, np.nan, np.nan],
   'AVIRIS_chl_a_25_09': [0.499, 0.542, 0.712, 0.849, 0.831, np.nan, 0.660]
}

df = pd.DataFrame(data)

#Testing different methods to deal with NaNs
# 1. Drop rows with missing values
SBC_dropped = df.dropna()
print("DataFrame with rows with missing values removed:")
print(SBC_dropped)

# 2. Impute missing values with the mean of the column
SBC_mean_imputed = df.fillna(df.mean())
print("\nDataFrame with missing values imputed with column mean:")
print(SBC_mean_imputed)

# 3. Interpolation (linear interpolation)
df_interpolated = df.interpolate(method='linear', limit_direction='forward')
print("\nDataFrame with missing values imputed by linear interpolation:")
print(SBC_interpolated)

#Testing normality of the above outputs/dataframes for SBC
# Function to apply Shapiro-Wilk test for each column in a DataFrame
```

```
def test_normality(df, description):
    print(f"\nNormality Test Results ({description}):")
    for col in df.columns[1:]:  # Skip 'Station' column
        stat, p = shapiro(df[col])
        print(f"{col}: W-statistic = {stat:.4f}, p-value = {p:.4f}")
        if p > 0.05:
            print(f"{col}: Data looks normally distributed (p > 0.05)\n")
        else:
            print(f"{col}: Data is not normally distributed (p <= 0.05)\n")

# Test normality for all three DataFrames
test_normality(SBC_dropped, "Rows with Missing Values Dropped")
test_normality(SBC_mean_imputed, "Mean Imputed Data")
test_normality(SBC_interpolated, "Interpolated Data")

#Since data is normally distributed, can apply Paired t-test as one of the statistical tests.
#Carry out statistical tests to determine accuracy of the imagery

# Function to calculate MAE, RMSE, and perform paired t-test
def calculate_accuracy(df, description):
    print(f"\nAccuracy Results ({description}):")

    # MAE, RMSE, Paired t-test for Sentinel-2 (19/09)
    print("\nSentinel-2 (19/09):")
    y_true = df['PnB_chl_a']
    y_pred_sentinel1 = df['Sentinel-2_chl_a_19_09']

    mae_sentinel1 = mean_absolute_error(y_true, y_pred_sentinel1)
    rmse_sentinel1 = np.sqrt(mean_squared_error(y_true, y_pred_sentinel1))
    t_stat_sentinel1, p_value_sentinel1 = stats.ttest_rel(y_true, y_pred_sentinel1)

    print(f"MAE = {mae_sentinel1:.4f}")
    print(f"RMSE = {rmse_sentinel1:.4f}")
    print(f"Paired t-test: t-stat = {t_stat_sentinel1:.4f}, p-value = {p_value_sentinel1:.4f}")
```

```
    # MAE, RMSE, Paired t-test for Sentinel-2 (24/09)
    print("\nSentinel-2 (24/09):")
    y_pred_sentinel2 = df['Sentinel-2_chl_a_24_09']

    mae_sentinel2 = mean_absolute_error(y_true, y_pred_sentinel2)
    rmse_sentinel2 = np.sqrt(mean_squared_error(y_true, y_pred_sentinel2))
    t_stat_sentinel2, p_value_sentinel2 = stats.ttest_rel(y_true, y_pred_sentinel2)

    print(f"MAE = {mae_sentinel2:.4f}")
    print(f"RMSE = {rmse_sentinel2:.4f}")
    print(f"Paired t-test: t-stat = {t_stat_sentinel2:.4f}, p-value = {p_value_sentinel2:.4f}")

    # MAE, RMSE, Paired t-test for AVIRIS (25/09)
    print("\nAVIRIS (25/09):")
    y_pred_aviris = df['AVIRIS_chl_a_25_09']

    mae_aviris = mean_absolute_error(y_true, y_pred_aviris)
    rmse_aviris = np.sqrt(mean_squared_error(y_true, y_pred_aviris))
    t_stat_aviris, p_value_aviris = stats.ttest_rel(y_true, y_pred_aviris)

    print(f"MAE = {mae_aviris:.4f}")
    print(f"RMSE = {rmse_aviris:.4f}")
    print(f"Paired t-test: t-stat = {t_stat_aviris:.4f}, p-value = {p_value_aviris:.4f}")

# Calculate accuracy for each DataFrame
calculate_accuracy(SBC_dropped, "Rows with Missing Values Dropped")
calculate_accuracy(SBC_mean_imputed, "Mean Imputed Data")
calculate_accuracy(SBC_interpolated, "Interpolated Data")
```