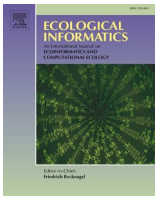
Ecological Informatics 94 (2026) 103626

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Laboratory-based hyperspectral reflectance analysis for phytoplankton species identification

R. Bentivogli [a,*], L. Pezzolesi [a,b,*], N. Caputo [b,c], B. Casarotto [d], S. Silvestri [a,b]

[a] Department of Biological, Geological and Environmental Sciences (BiGeA), University of Bologna, P.za Porta S. Donato 1, 40126 Bologna (BO), Italy
[b] Interdepartmental Centre for Research in Environmental Sciences (CIRSA), University of Bologna, Via Sant'Alberto 163, 48123 Ravenna (RA), Italy
[c] Marine and Freshwater Research Centre, Atlantic Technological University, Dublin Road, H91 T8NW Galway, Ireland.
[d] Department of Geosciences, University of Padua, Via Gradenigo 6, 35131 Padova (PD), Italy

ARTICLE INFO

Keywords:
Phytoplankton
Spectral analysis
Hyperspectral sensor
Microalgal pigments
Algal community
HABs monitoring

ABSTRACT

Monitoring phytoplankton communities is essential for assessing ecosystem health and detecting harmful algal blooms (HABs). Hyperspectral imaging has emerged as a promising tool to discriminate among microalgal species based on their unique reflectance signatures. This study presents a laboratory spectral analysis of five phytoplankton species, including bloom-forming and toxin-producing taxa common in coastal waters. Reflectance spectra were measured at multiple cell concentrations and analyzed using two normalization approaches, second- and fourth-derivative transformations, and dimensionality reduction techniques including principal component analysis (PCA) and linear discriminant analysis (LDA).

Our results demonstrate that specific spectral features, particularly in the 470–500 nm and 620–680 nm ranges, enable species-level discrimination. PCA and LDA effectively enhanced separability by reducing spectral redundancy and emphasizing class features. We further applied linear spectral unmixing (LSU) to estimate fractional species abundances in synthetic mixtures. LSU performed well in simple mixtures but revealed limitations in complex communities, where nonlinear effects and spectral similarity reduced accuracy.

Beyond classification, LSU enables quantitative assessment of species contributions, providing a valuable complement to PCA and LDA for ecological interpretation and bloom dynamics investigation. This integrated approach lays the foundation for future development of operational tools that combine spectral unmixing and machine learning for automated HAB detection. The combined use of hyperspectral reflectance data and computational methods supports scalable, real-time monitoring of phytoplankton diversity and abundance, with strong potential for deployment in early-warning systems and coastal observatories.

1. Introduction

Phytoplanktonic organisms play a critical role in global carbon cycling and nutrient dynamics, serving as primary producers and forming the base of the aquatic food web in both marine and freshwater ecosystems (Kudela et al., 2017; Marampouti et al., 2021). As responsible for half of the Earth's primary production, phytoplankton directly influences the abundance and diversity of marine and freshwater organisms (Maberly and Gontero, 2022) and its dynamics are susceptible to environmental changes, which can induce structural and functional shifts at the cellular level (Dashkova et al., 2022; Stelmakh et al., 2023). These shifts impact ecological interactions and affect water quality, as phytoplankton can alter parameters such as nutrient availability, dissolved oxygen levels, and turbidity.

It is well-known that certain phytoplankton species produce toxins that are detrimental to aquatic life, human health, and the economy, underscoring the importance of monitoring algal communities, especially in ecosystems vulnerable to harmful algal blooms (HABs) and highlighting the need for efficient and timely detection methods (Davidson et al., 2014; Dorantes-Aranda, 2023). Quantifying phytoplankton biomass, community composition and cell-size distribution is also critical for documenting the onset and the evolution of HABs and their associated ecological and public health impacts (Pahlevan et al., 2021).

* Corresponding authors.
E-mail addresses: riccardo.bentivogli4@unibo.it (R. Bentivogli), laura.pezzolesi@unibo.it (L. Pezzolesi), nicole.caputo@research.atu.ie (N. Caputo), bruno.casarotto@phd.unipd.it (B. Casarotto), sonia.silvestri5@unibo.it (S. Silvestri).

https://doi.org/10.1016/j.ecoinf.2026.103626
Received 5 June 2025; Received in revised form 20 January 2026; Accepted 20 January 2026
Available online 21 January 2026

Traditional monitoring methods rely on labor-intensive microscopic analysis, which requires significant expertise and time, making them impractical for large-scale, real-time applications (Penna et al., 2017; Santi et al., 2021; Wang et al., 2022).

On the other hand, satellite-based ocean color remote sensing has been widely used in recent years to estimate chlorophyll *a* concentration as a proxy for phytoplankton biomass across broad spatial scales (Cuartero et al., 2023; Feng et al., 2015; Niroumand-Jadidi et al., 2021). Aquatic remote sensing has long been proven to be an efficient means for detecting harmful algal blooms (HABs) and assessing phytoplankton biodiversity (Pahlevan et al., 2021; Sathyendranath, 2014). However, most traditional satellite-based approaches have relied on band indices or empirical algorithms derived from multispectral sensors to estimate chlorophyll *a*, the primary proxy for phytoplankton biomass (Bresciani, 2011; Gower et al., 2008). While these methods have been successful for large scale monitoring, their limited number of spectral bands often restricts the ability to resolve community composition or discriminate among phytoplankton taxa, particularly when species compositions are diverse or mixed (He et al., 2020; Rocchini et al., 2022). A complete characterization of phytoplankton biodiversity requires high-fidelity retrievals of pigment-specific features that can only be achieved through high-resolution optical data such as those provided by hyperspectral sensors (Pahlevan et al., 2021).

In aquatic applications, imaging spectrometers typically provide continuous spectral coverage in the visible and near-infrared (VNIR), generally spanning between 400 and 1000 nm (Dierssen et al., 2021; Jaywant and Arif, 2024). The spectral range of water-leaving remote-sensing reflectance is primarily constrained by pure-water absorption, making the VIS–VNIR region the most informative domain for phytoplankton optical characterization (Hieronymi et al., 2023). Compared to multispectral sensors that sample only a limited number of broad, discontinuous bands, hyperspectral data provide hundreds of narrow, adjacent bands that capture subtle spectral features and can distinguish minute differences in spectral signatures relevant to species discrimination (Braga et al., 2022; Fabbretto et al., 2024; Jaywant and Arif, 2024).

Recent advancements in remote sensing technologies, particularly hyperspectral imaging, have emerged as promising alternatives for coastal water quality studies, which are typically optically complex waters. Many existing applications have focused on suspended sediment and turbidity retrieval (Zhou et al., 2017), whereas phytoplankton-related applications typically involve more complex spectral analyses. Hyperspectral imaging enables the rapid and accurate assessment of algal communities by taking advantage of the unique optical signatures associated with different phytoplankton species and their specific pigment contents (Bi et al., 2019; Solovchenko, 2023). Pigments such as chlorophylls, carotenoids and phycobilins vary significantly among microalgal species, giving them distinct optical characteristics (Table 1). Chlorophylls, for example, primarily absorb light in the blue and red regions of the visible spectrum, while carotenoids absorb blue-green light and phycobilins absorb green-yellow light.

This diversity of pigments allows microalgae to utilize different parts of the light spectrum for photosynthesis, influencing how they reflect, absorb, and transmit light. Laboratory-based studies using hyperspectral imaging have demonstrated the ability to capture these differences in reflectance spectra among various microalgal species (Salmi et al., 2021, 2022; Soja-Woźniak et al., 2018), providing essential reference data for developing algorithms that can be applied to field and remote sensing observations. The spectral response of phytoplankton is influenced by the cellular characteristics such as size, shape, and pigment composition, as well as by environmental factors like nutrient availability and light conditions (Bricaud et al., 2004; Mao et al., 2010; Nair et al., 2008). Moreover, in environmental monitoring, Case 2 waters present challenges related to the presence of various optically active components, such as suspended sediments, colored dissolved organic matter (CDOM), and other particulates, which could mask or alter phytoplankton optical signals, complicating the analysis due to high spectral overlap and redundancy information (Dierssen et al., 2021; Vaddi et al., 2024).

The integration of advanced machine learning techniques into hyperspectral data analysis pipelines has enhanced the capacity to differentiate phytoplankton taxa in optically complex waters characterized by mixed communities (Fournier et al., 2024; Pyo et al., 2019; Xu et al., 2020).

Beyond traditional machine-learning classifiers, deep-learning architectures (e.g., CNN-based models) have been widely explored for hyperspectral image classification by learning discriminative spatio-spectral features (Lumini and Nanni, 2019; Yang et al., 2023). However, deep networks can be limited by the availability of training data and may overfit when training samples are insufficient, motivating the use of transfer learning and fine-tuning strategies. Recurrent models have also

Table 1
List of the main pigments of various microalgae groups (modified from source: Roy et al., 2011).

Pigment	Green Algae	Diatom	Dinoflagellate	Cyanobacteria	Ochrophyta	Absorption Range
Chlorophylls						
Chl a	•	•	•	•	•	430-470 nm; 665-680 nm
Chl b	•					450-470 nm; 640-660 nm
Chl c1		•			•	440-470 nm; 620-640 nm
Chl c2		•	•		•	450-480 nm; 620-640 nm
Chl c3		•				450-480 nm; 620-630 nm
Carotenoids						
Antheraxanthin	•					420-480 nm
B-carotene	•	•	•			420-500 nm
Diadinoxanthin		•	•		•	450-480 nm
Diatoxanthin		•	•			450-480 nm
Dinoxanthin			•			410-480 nm
Fucoxanthin		•			•	450-470 nm
Lutein	•					470-500 nm
Myxoxanthophyll				•		420-490 nm
Peridinin			•			480-485 nm
Violaxanthin	•				•	470-500 nm
Zeaxanthin				•		420-490 nm
Phycobilins						
Allophycocyanin				•		610-650 nm
Phycocyanin				•		650-660 nm
Phycoerythrin				•		540-570 nm

been proposed to exploit the sequential nature of contiguous spectral bands, such as bidirectional GRU-based networks for hyperspectral image classification (Wu et al., 2023). In aquatic remote-sensing applications, deep-learning models (including hybrid CNN–LSTM architectures) have also been used with satellite products to forecast chlorophyll dynamics and algal blooms (Ding and Li, 2024; Yao et al., 2023).

Although machine learning models are promising for enhancing the interpretability and classification of spectral data, a disconnect remains between laboratory-based spectral libraries and field-scale classification models. To date, most laboratory-based optical studies have focused on freshwater organisms, particularly cyanobacteria (Adejimi et al., 2023; Fournier et al., 2024; Malhotra and Örmeci, 2023; Pokrzywinski et al., 2021), due to the unique pigments that make them easier to distinguish from other phytoplankton taxa. However, given the high taxonomic diversity of phytoplankton in oceanic and coastal waters, it is essential to spectrally differentiate the most common phytoplankton groups using hyperspectral data from advanced optical sensors (Xi et al., 2015). Despite this need, relatively few studies have investigated the optical differences between species belonging to the same phytoplankton group (Aparicio-Rizzo et al., 2025; Bi et al., 2019; Grunert et al., 2025; Xi et al., 2015; Zhang et al., 2024). This gap is particularly relevant because, even within the same group, both HAB-forming organisms and non-HAB species can coexist. Distinguishing among these organisms would be crucial for improving species-specific monitoring and management of algal communities in natural environments.

This study explores the application of a hyperspectral sensor for laboratory analysis of various microalgae by measuring and analyzing the reflectance profiles of five monospecific phytoplankton cultures belonging to different taxonomic groups. The selected species are commonly found in the Mediterranean Sea, with some of them also being responsible for harmful algal blooms in the Northern Adriatic (Čanković et al., 2022; Valbi et al., 2019; Zingone et al., 2021). By systematically examining the spectral responses of both monospecific and mixed algal cultures at cell densities which mimic environmental conditions, this research aims to better understand the reflectance properties able to distinguish the different phytoplankton compositions. The analysis focuses on identifying optical features and the wavelengths that are most sensitive for distinguishing algal species, to select the most effective ones for accurate species differentiation. This study serves as a laboratory exercise that could lay the groundwork for a deeper understanding of the potential to work with reflectance data, to apply these insights to satellite-based studies. These advancements are crucial for the timely detection of HABs. As research progresses, the application of hyperspectral remote sensing combined with advanced data analysis holds great potential for transforming our capacity to monitor and manage aquatic coastal ecosystems in current conditions as well as in climate change scenarios, providing a strong foundation for further field research and the development of integrated phytoplankton monitoring strategies.

2. Materials and methods

2.1. Phytoplankton cultures and biomass quantification

Five different species of microalgae belonging to different algal groups were selected and cultured: *Alexandrium minutum* (strain AMC1801, HABs and PSP toxin producer), and *Heterocapsa* sp. (strain HESG2001, HABs producer) for the Dinophyceae class, *Chattonella* sp. (strain CHSG2201, HABs producer) as a representative of the Raphydophyceae class, *Phaeodactylum tricornutum* (strain PTN0301, bloom-forming species) for the Bacillariophyceae class, and *Synechococcus* sp. (strain CCMP1631, abundant in coastal areas playing a major role in both carbon cycling and trophic networks) for the Cyanophyceae class. Each species was cultured in f/2 (+Si for the diatom *P. tricornutum*) culture medium (Guillard and Ryther, 1962) at a salinity of 35 psu under stable conditions of temperature (20 ± 1 °C) and light (120 ± 10 μE m^{-2} s^{-1}) with a photoperiod of 16:8 h (light:dark). Cultures were prepared at 6 different cell densities for each algal species (see table S1) using dilutions of the mother cultures at known cell concentrations. For cell counting, the Utermöhl method (Utermöhl, 1931) was employed using an inverted microscope at 320× magnification (Zeiss Axiovert 100). Because of the small size of *Synechococcus* sp., for this species cell counting was performed using a Bürker counting chamber. Biovolume calculation for each algal species followed Hillebrand et al. (1999). Chlorophyll *a* concentration was determined spectrophotometrically (UV/VIS, JASCO V-650, Tokyo, Japan0) at 750 and 665 nm (APAT, IRSA-CNR, 2003) using 90% acetone method and GF/C filters (Whatman, 47 mm); the determination of phaeopigments was performed through acidification of the extract with HCl (0.66 M).

2.2. Hyperspectral image acquisition system

The hyperspectral sensor used was a Headwall Photonics Nano-Hyperspec, a push-broom camera ranging in the VNIR (400 to 1000 nm) with 270 spectral bands. This high spectral sampling is advantageous over multispectral imagery because narrow, contiguous bands improve sensitivity to subtle spectral differences relevant for aquatic ecosystem characterization. The laboratory setup followed the general reflectance-mode configuration described by Lodhi et al. (2019), consisting of a hyperspectral camera, illumination source, and translation stage aligned to collect spatial and spectral information line by line. Our custom setup included a motorized stage (20 × 20 cm) to allow controlled sample translation, and a 150 W halogen lamp positioned at approximately 45° to the water surface, 25 cm above the surface to ensure uniform illumination while minimizing specular reflections.

A Spectralon 99% reflectance panel (Labsphere, Inc.) was placed beneath the transparent sample container to provide a known reference for radiance-to-reflectance conversion and to maintain consistent optical conditions across measurements. Before measuring algal samples, a baseline spectrum of pure water was acquired to account for background effects. The setup used in this study is shown in Fig. 1.

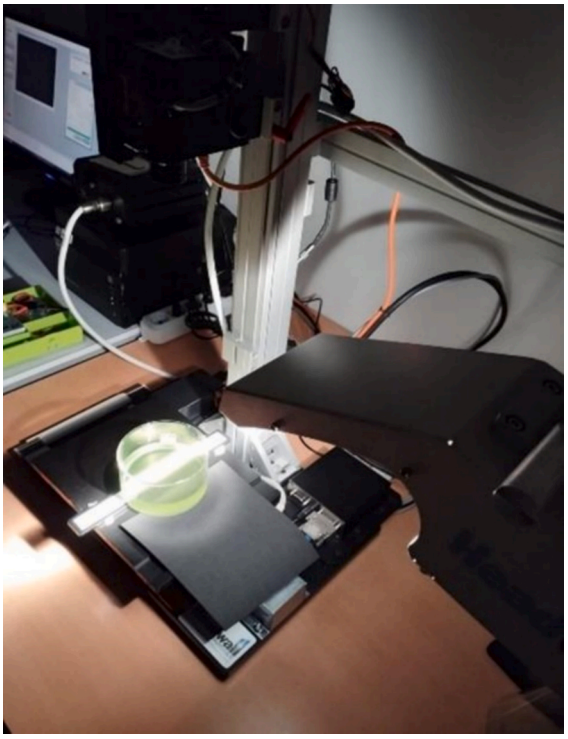


Fig. 1. Container and the setup used for the measurements.

2.3. Sample preparation

For each algal species, six sub-samples were prepared from serial dilutions of the stock culture to evaluate six different cellular concentrations following the scheme outlined in the supplementary material (Table S1, Fig. S1). The final concentrations of the stock cultures were defined with the aim of simulating realistic scenarios of cellular bloom concentrations in the environment (Pistocchi et al., 2012). From each sub-culture, a volume of 200 mL was placed in a flat glass container that was then used for performing the measurements with the hyperspectral sensor (Fig. 1).

The white reference panel, together with the container, was placed beneath the sensor to provide a uniform and known bottom reflectance and was used to convert the raw radiometric data into reflectance values.

2.4. Data collection and spectral analysis

Reflectance measurements were obtained for each species at various cellular concentrations. Images were pre-processed before analyzing spectral features to ensure high-quality data. For each image, a Region of Interest (ROI) was selected to narrow down to a subset of 500 × 500 pixels, and the reflectance values of the selected pixels were averaged. Additionally, wavelengths above 790 nm, considered non-informative for this study, were filtered out.

Following pre-processing, the reflectance spectra were analyzed and both raw reflectance spectra and normalized reflectance spectra were examined. Spectral normalization was performed in two ways: first, by biovolume, which adjusted for the average cell volume of each species to better compare reflectance efficiency per unit of biomass. The normalization of reflectance by biovolume was performed by dividing the reflectance values of each pixel by the total biovolume of the corresponding species, expressed in μm^3 per 200 mL. This approach allows for the comparison of reflectance among species with different cell sizes, reducing the influence of biomass differences on spectral reflectance. The second method was normalization by maximum reflectance values, which emphasized the shape of each spectrum, making the spectral signatures of different species more directly comparable. This type of normalization is particularly useful in remote sensing and environmental monitoring, where normalized reflectance data are better suited for machine learning models and statistical analyses.

To further enhance the analysis of spectral features, the second derivative of the reflectance spectra was calculated. Before computing the second derivative, data were smoothed using the Savitzky-Golay filter to reduce noise and enhance the quality of the spectral data (Savitzky and Golay, 1964). This data smoothing technique uses a weighted moving average to smooth the data while preserving critical features such as peaks and rapid variations. It effectively reduces noise without distorting the spectral characteristics, making it widely used in spectroscopy. For this study, the smoothing parameters were set to a window size of sixteen and a third-order polynomial fit. These settings were carefully selected after assessing the hyperspectral data to maintain the characteristic details of the reflectance spectrum. Here, the window size indicates the number of data points included in the smoothing process, while the polynomial order refers to the degree of the polynomial used for fitting the data within each window.

The second derivative analysis highlighted specific wavelengths associated with pigment absorption and is widely recognized as an effective tool for assessing microalgal pigment features (Aguirre-Gómez et al., 2001; Méléder et al., 2013). This derivative approach, combined with biovolume and maximum-value normalization, provided a heterogeneous view of each species' spectral signature.

Additionally, the second derivative can help reduce low-frequency noise, correct baseline shifts caused by instrumental variations, and improve spectral resolution by separating overlapping peaks (Méléder et al., 2013). This multi-step preprocessing ensured that the reflectance data were optimized for subsequent analyses, allowing for a more accurate characterization of the algal species. Moreover, with the goal of studying whether the spectral characteristics of the species remained consistent as concentrations decreased, normalized data were further processed using the Savitsky-Golay smoothing method, followed by calculation of the fourth derivative (see supplementary Fig. S2). These transformations were then combined to calculate the Similarity Index (SI), indicated in Eq. (1) as defined by Millie et al. (1997):

$$SI = 1 - \frac{2\arccos(A)}{\pi}$$

where A represents the cosine of the angle between two vectors derived from the fourth derivative of reflectance spectra, smoothed with a Savitsky-Golay filter. When SI equals 1, it indicates complete similarity between the reference and derived spectra, whereas a SI of 0 denotes no similarity. To apply the SI to our data, we selected a “reference spectrum” corresponding to the highest biomass concentration of each phytoplankton species (C1), while the “derived spectra” were obtained from lower concentrations (C2-C6). The SI quantifies the similarity between the reference and derived spectra and assesses how characteristic spectral features change with decreasing cellular concentration. Before computing the SI, spectral data were normalized using a single-step maximum value normalization, rather than first normalizing by biovolume and then by maximum value. Preliminary analyses showed that sequential normalization did not significantly affect SI results compared to using only the maximum value normalization. While biovolume normalization aimed to adjust for biomass differences, maximum value normalization was found to be more effective in highlighting spectral shape differences among species. This method allowed for a clearer comparison of spectral features, focusing on the relative shape of the reflectance spectra, which was critical for differentiating between species in our dataset.

2.5. Dimensionality reduction and feature extraction

Hyperspectral data contain many contiguous and highly correlated spectral bands, resulting in a substantial data volume that can be challenging to manage and analyze effectively. To reduce data dimensionality and identify the most informative spectral bands, Principal Component Analysis (PCA) was applied to the reflectance spectra. Prior to PCA, the reflectance data were standardized using z-score normalization (mean = 0, variance = 1) to ensure that all variables contributed equally to the analysis. The dataset was constructed by randomly selecting 500 pixels (100 pixels per species) to ensure balanced representation across the species. An initial PCA was conducted on the standardized reflectance spectra without additional species normalization or biovolume correction. This was performed to explore the primary sources of variation in the dataset. However, the results showed a strong linear alignment of samples along PC1 and PC2, suggesting that the dominant source of variance was overall intensity differences rather than species-specific spectral features. To refine the analysis and emphasize spectral differences, a second PCA was performed after normalizing reflectance values within each species. This normalization adjusted for interspecies variations in absolute reflectance while preserving spectral shape differences. Furthermore, biovolume information was incorporated as a variable to account for differences in sample concentration and ensure that species with stronger reflectance signals were not represented disproportionately. This refined approach reflected intrinsic spectral characteristics rather than intensity-driven variance.

To evaluate the discriminatory power of the spectra, we also applied Linear Discriminant Analysis (LDA). After z-score standardization, the spectra were reduced to the first three principal components. These PC scores were then used as input to LDA. Whereas PCA maximizes total variance, LDA focuses on maximizing class separability between

predefined classes (i.e., species) and was used on the standardized raw spectra therefore tests how well the data can distinguish the taxa based on their spectral shapes, even in the presence of overall intensity differences.

2.6. Data decomposition and spectral unmixing

In addition to analyzing the reflectance spectra of individual species at different cellular concentrations, we applied a spectral unmixing technique to the reflectance data from algal mixtures. To isolate the spectral signature of each algal species (endmembers), the water spectral contribution was removed using a volumetric correction approach. Specifically, for each individual algae, the total reflectance (R_{total}) was corrected by subtracting the reflectance contribution of water (R_{water}) weighted by the water volume (f_{water}) relative to the algal biovolume ($f_{species}$), that was previously calculated, according to the Eq. (2).

$$R_{species} = \frac{R_{total} - R_{water} \bullet f_{water}}{f_{species}}$$

where (R_{water}) represents the spectral reflectance of pure water, measured independently previously. A similar procedure was applied to the mixed algal samples, where the contribution of water was removed by considering the specific biovolumes of each algal species within the mixed solution. This produced reflectance spectra corrected for water contributions, for both individual algae (endmembers) and the algal mixture, ensuring that subsequent spectral analyses were unaffected by water. The mixtures contained a total volume of 200 mL, with specific proportions of each species (Table 2) calculated based on biovolume to achieve controlled ratios (microalgal biovolumes contained in the single mixtures are reported in Table S2).

Mix samples were prepared to include two, three and four representative species from each algal group in order to have an increasing complexity. In mix 3, *A. minutum* was selected (instead of *Heterocapsa* sp.) as the representative dinoflagellate due to its relevance as a toxic species of high monitoring interest, but also by its ecological importance, as it is among the most commonly blooming dinoflagellates in Mediterranean coastal waters, avoiding spectral redundancy and ensuring a balanced multispecies assemblage. These controlled ratios, along with each species' known biovolume, allowed to assess species contributions within the mixed samples. Linear Spectral Unmixing (LSU) was subsequently applied to estimate the fractional abundance of each algal species within the mixed sample. LSU assumes the observed spectral reflectance of each pixel to be a linear combination of identified endmembers, according to equation Eq. (3).

$$P_i = \sum_{J=1}^{n} \left(R_{ij} \cdot F_j\right) + E_i$$

where:

- P_i is the observed spectral reflectance of the i^{th} spectral band for a given pixel,
- R_{ij} is the known spectral reflectance of the j^{th} endmember in the i^{th} spectral band,
- F_j is the fractional abundance of the j^{th} endmember in the pixel,
- E_i is the error term representing unmodeled reflectance and noise.

Table 2
Species composing the mixtures and their respective ratio by biovolume.

Species	Mix 1	Mix 2	Mix 3
A. minutum	1		1
Chattonella sp.			1
Heterocapsa sp.	1	1	
P. tricornutum		1	1
Synechococcus sp.		1	1

To ensure physically meaningful and interpretable results, two constraints were applied during the LSU optimization: the non-negativity constraint Fj $\geq$ 0, which ensures that the estimated fractional abundances of all endmembers are non-negative, and the $\sum_{j=1}^{n} F_j = 1$, that guarantees that the fractional abundances sum exactly to unity. The optimization problem was solved numerically using the OSQP solver available in the CVXPY library available in Python. To assess the accuracy of the spectral unmixing results, we evaluated the reconstructed spectra against the observed mixed spectra using two key metrics:

- Root Mean Square Error (RMSE): RMSE was computed to quantify the overall reconstruction error as indicated in Eq. (4).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(R_{observed,i} - R_{reconstructed,i}\right)^2}$$

Where $R_{observed,i}$ is the observed reflectance at the i^{th} wavelength and $R_{reconstructed,i}$ is the corresponding reconstructed reflectance obtained from the LSU model. RMSE was computed to assess the reconstruction accuracy of LSU. Since reflectance values are dimensionless, RMSE is reported as a unitless measure, directly representing the deviation between observed and reconstructed reflectance. Lower RMSE values indicate better spectral reconstruction.

- Spectral Angle Mapper (SAM): SAM was used to evaluate the similarity between the observed and reconstructed spectra by measuring the spectral angle in multi-dimensional space, following Eq. (5).

$$SAM = \cos^{-1}\left(\frac{R_{observed} \bullet R_{reconstructed}}{\left(\sqrt{\sum_{i=1}^{n} R_{observed,i}^2}\right) \bullet \left(\sqrt{\sum_{i=1}^{n} R_{reconstructed,i}^2}\right)}\right)$$

where the dot product is taken over all wavelengths. The SAM metric provides an angular difference (in degrees), with lower values indicating higher spectral similarity. These metrics provide quantitative validation of the LSU performance. To ensure robust validation, 100 pixels were selected for each endmember and 100 pixels for each mixture. RMSE and SAM were computed by comparing the observed and reconstructed reflectance spectra across these 100 pixels per condition. This approach accounts for variability across different pixels and ensures a statistically sound evaluation of the LSU performance. Additionally, we compared the mean reconstructed spectrum with the mean observed mixed spectrum through visual inspection to further assess the accuracy of LSU. To quantify LSU performance in estimating species abundances, we calculated the mean absolute abundance error (MAAE), defined as the mean absolute difference between LSU-derived fractional abundances and the actual fractions measured from biovolume for each mixture:

$$MAAE = \frac{1}{n} \sum_{j=1}^{n} \left|F_{j,\text{estimated}} - F_{j,\text{true}}\right|$$

where *n*is the number of algal species present in a given mixture. MAAE directly measures the accuracy of fractional abundance retrieval and complements RMSE and SAM by evaluating errors in species composition rather than spectral similarity. This metric was also used to assess the performance drop of LSU across mixtures of increasing complexity, allowing us to quantify how estimation accuracy degrades when spectral overlap among species becomes stronger.

2.7. Software and tools

Images pre-processing steps such as ROI extraction, spectral profile analysis, and pixel averaging were performed using ENVI® Classic version 5.3 (Exelis Visual Information Solutions, now NV5 Geospatial, Boulder, Colorado). ENVI® is a registered trademark of NV5 Global, Inc. The spectral analyses were performed using Python (Python Software Foundation, version 3.11.4). The SciPy library (version 1.11.1) was used for data smoothing and filtering, Matplotlib (version 3.7.1) for creating plots and visualizations, and Scikit-learn (version 1.3.0) for Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Spectral unmixing was performed using the cvxpy library (version 1.6.0) for constrained optimization and NumPy library (version 2.2.3) for matrix operations.

3. Results

3.1. Microalgal culture characterization

Each algal culture at the maximum concentration (C1) was characterized in terms of cellular counts, cell biovolume measurements, and chlorophyll *a* content; results are presented in Table 3.

Measurements of biovolume, cell counts, and chlorophyll *a* content were conducted for all the algal species using cultures in the stationary growth phase. *A. minutum* showed a mean cell concentration of $2.9 \times 10^5 \pm 9.1 \times 10^4$ cells L^{-1} and a mean biovolume of 2920 ± 766 μm^3, while *Chattonella* sp. recorded lower values ($9.3 \times 10^4 \pm 5.4 \times 10^3$ cells L^{-1},1691 ± 437 μm^3). *Synechococcus* sp. had the highest cell density (3.2×10^7 cells L^{-1}), but the smallest biovolume (26 ± 2 μm^3). *P. tricornutum* exhibited a mean biovolume of 100 ± 55 μm^3, while *Heterocapsa* sp. showed intermediate values ($9.3 \times 10^5 \pm 1.6 \times 10^5$ cells L^{-1}, 129 ± 37 μm^3). The relatively high standard deviation observed for *Phaeodactylum tricornutum* cell counts (Table 3) likely reflects the heterogeneous spatial distribution of cells during microscopic enumeration. This species is characterized by very small cells with a marked morphological variability, occurring in multiple morphotypes (oval, fusiform, and triradiate), which can lead to uneven dispersion within the counting chamber. As a result, slight differences in cell settling or local concentration across transects can produce substantial variability among replicate counts.

The chlorophyll *a* content of the cultures varied among species, reflecting known differences in pigment composition and cell size. *P. tricornutum* showed the highest chlorophyll *a* concentration (1.3×10^3 mg m^{-3}), whereas *Chattonella* sp. had the lowest (2.8×10^2 mg m^{-3}). Although *Synechococcus* sp. exhibited a high total cell density, its total chlorophyll *a* concentration was moderate (6.8×10^2 mg m^{-3}), likely reflecting its very small cell size and inherently low pigment content per unit cell volume typical of picocyanobacteria. This observation is consistent with the physiological characteristics of *Synechococcus* and similar picoplanktonic taxa, which generally have less chlorophyll *a* per cell compared with larger eukaryotic algae.

Table 3
Measurements of cell density, biovolume and chlorophyll *a* of the tested species at the maximum concentration (C1).

	Cell Density (cells L^{-1})		Biovolume (μm^3)		Chlorophyll *a* (mg m^{-3})	
	mean	st.dev.	mean	st.dev.	mean	st.dev.
A. minutum	2.9×10^5	9.1×10^4	2.9×10^3	7.7×10^2	4.0×10^2	1.9×10^1
Chattonella sp.	9.3×10^4	5.4×10^3	1.7×10^3	4.4×10^2	2.8×10^2	4.8×10^1
Heterocapsa sp.	9.3×10^5	1.6×10^5	1.3×10^2	3.7×10^1	4.6×10^2	1.4×10^1
P. tricornutum	4.3×10^6	2.7×10^6	1.0×10^2	5.5×10^1	1.3×10^3	6.6×10^1
Synechococcus sp.	3.2×10^7		2.6×10^1	2.55×10^0	6.8×10^2	2.7×10^1

3.2. Spectral data analysis

First, the spectral signatures of the five tested species were analyzed across the various concentrations (Fig. 2). As expected, curves corresponding to the C1 concentration showed the lowest reflectance, that generally increased as the cell concentration decreased. A general observation for all species was that, as cell concentrations decreased, the absorption and reflection peaks became progressively less distinct, showing a tendency to flatten.

The reflectance scale varied among species, ranging from *Synechococcus* sp., which showed very low reflectance values (≈0.2) at lower wavelengths, to *Chattonella* sp., which started with significantly higher values (≈0.6) at the lowest wavelength. Another notable characteristic was that the positions of peaks shifted by a few nanometers toward shorter wavelengths as concentrations decreased; this is particularly noticeable at 709–711 nm, where the peak gradually shifted toward shorter wavelengths as the concentration decreased. To evaluate the differences among spectra of various species, the spectral signatures of the highest concentration (C1) were analyzed following two distinct normalization approaches to enhance interspecies comparability. Additionally, the spectral signature of the culture medium (Fig. 2F), composed of water and nutrients, showed a relatively flat reflectance curve across all wavelengths, with a slight decrease in reflectance beyond 700 nm. This is consistent with the expected optical behavior of water, which exhibits low variability in the visible spectrum and a characteristic drop in reflectance in the near-infrared region. The medium's spectral response served as a reference to separate the contributions of microalgae to the observed reflectance curves. Normalization by biovolume indicated that *Heterocapsa* sp. exhibits the highest reflectance across most wavelengths (Fig. 3A), followed by *Chattonella* sp., which showed a flattened shape. Conversely, *A. minutum, P. tricornutum,* and *Synechococcus* sp. displayed a consistently lower reflectance across the spectrum. To further investigate the spectral shape independently of absolute reflectance intensities, a secondary normalization was applied, dividing each spectrum by its maximum reflectance value (Fig. 3B).

This approach emphasized the unique spectral shape of each species, facilitating comparisons based on relative spectral features rather than intensity. Following shape normalization, all species showed strong absorption in the blue region between 470 and 500 nm, while a high reflectance is observed after 700 nm in the red-edge region. However, distinctive peaks and troughs emerged for the various species. *Chattonella* sp. displayed a relatively uniform spectral shape with a little peak at 650 nm, while *Heterocapsa* sp. and *A. minutum* exhibited overlapping spectral profiles in the visible region, but the first one showed slightly lower reflectance at shorter wavelengths (<550 nm). They both had a distinct peak at 607 nm before flattening until 656 nm where the spectra started to decrease. Conversely, *P. tricornutum* and *Synechococcus* sp. exhibited the most distinct spectral shapes compared to the other species, with unique reflectance peaks and troughs. *P. tricornutum* showed two reflectance peaks, one around 605 nm similar to those observed in *Heterocapsa* sp. and *A. minutum*, and the other at 650 nm; it also exhibited stronger absorption between the two peaks (around 635 nm) and a distinct shape between 470 and 500 nm. *Synechococcus* sp., on the other hand, displayed a unique spectrum with a reflectance peak around 550 nm, which is not observed in any of the other species. Reflectance in the visible range reached low values around 640 nm, where a strong absorption is visible, and then rose again at 660 nm before dropping near 680 nm, where low values are also observed in the other species. Using the data normalized to maximum values, the curves of the different species were analyzed in terms of relative differences (Fig. 4).

It can be observed that *Synechococcus* sp. indeed displayed the highest relative difference compared to *A. minutum*, *Heterocapsa* sp., and

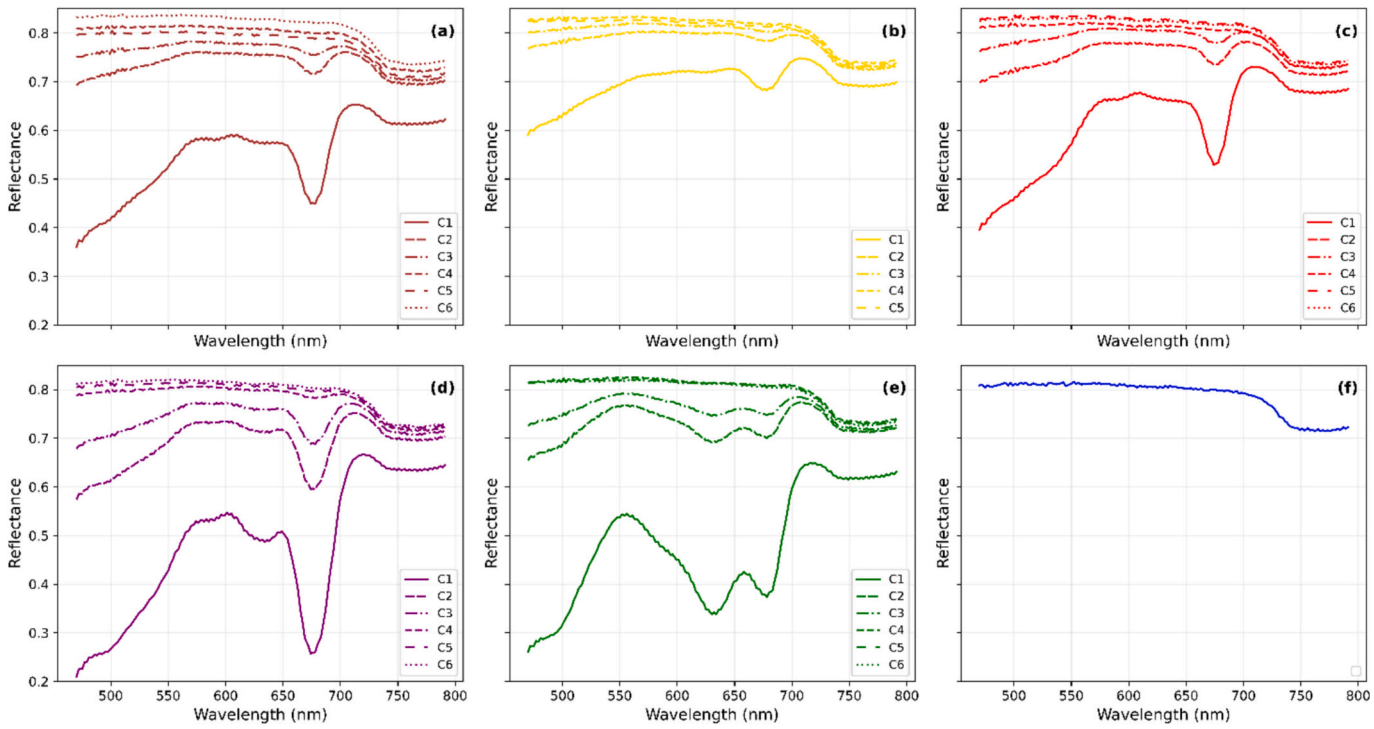


Fig. 2. Reflectance spectra of the five phytoplankton species at various cellular concentrations (C1–C6) (A-E). | (a) *Alexandrium minutum*, (b) *Chattonella* sp., (c) *Heterocapsa* sp., (d) *Phaeodactylum tricornutum*, (e) *Synechococcus* sp., and (f) the culture medium ("water + nutrients") used for growing the microalgae.

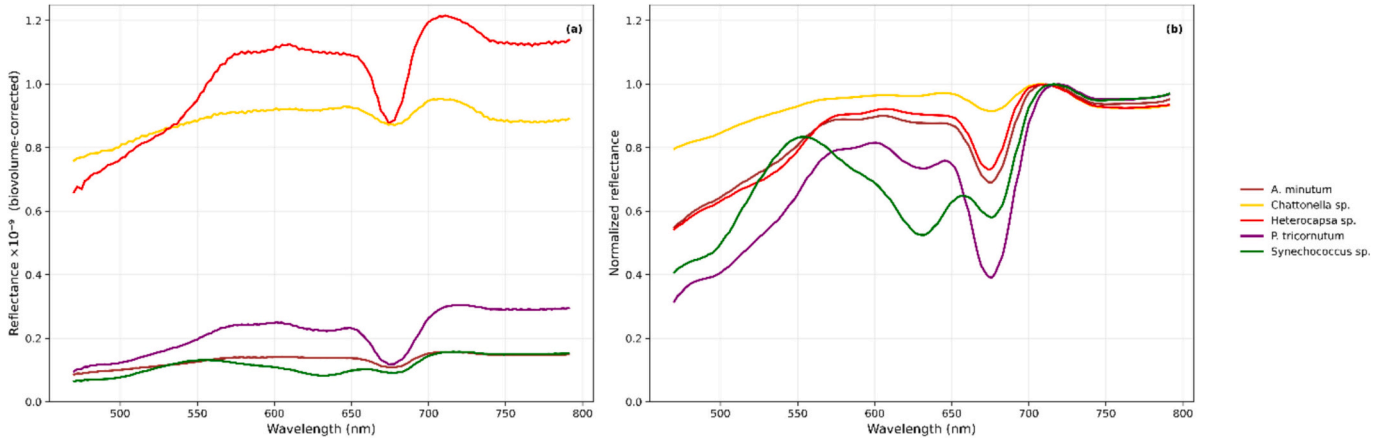


Fig. 3. Reflectance spectra of five phytoplankton species after preprocessing. | (a) Spectra corrected for cell biovolume (μm^3 per 200 mL) and displayed as reflectance $\times 10^{-9}$, illustrating the relative efficiency of each species per unit biomass. (b) Spectra normalized to their maximum reflectance and smoothed with a Savitzky–Golay filter (window = 16 bands, polynomial = 3), emphasizing differences in spectral shape.

Chattonella sp. (Fig. 4E). The wavelengths where this difference was mostly pronounced are between 480 and 490 nm, 625–630 nm, and 680 nm. On the other hand, *P. tricornutum* showed a similar pattern to *Synechococcus* sp., especially at 570 nm and 660 nm where the relative difference reached the lowest values (Fig. 4D). *Heterocapsa* sp. and *A. minutum* show very low differences between each other and relatively low differences when compared with *Chattonella* sp., therefore spectra in Fig. 4A and Fig. 4C looked very similar.

Since all species exhibited maximum normalized reflectance values around 709 nm, the relative differences between them become lower at higher wavelengths. To further assess spectral differences between species and observe their behavior across different concentrations, the second derivative of the spectra was calculated for each algal species (Fig. 5), using a uniform y-axis scale to facilitate direct comparison of spectral intensity across all species.

In general, it can be observed that the main peaks are concentrated between 630 and 700 nm for all the tested samples, which suggests potential spectral markers for differentiating among species. The most distinctive peak for all the species was at around 680 nm. *A. minutum* displayed a relatively flat curve, especially between 500 and 550 nm (Fig. 5A), similar to *Heterocapsa* sp., which exhibited a noticeable peak near 700 nm (Fig. 5C), but with a slightly different pattern in magnitude and shape. In contrast, *Chattonella* sp. displayed some of the lowest second derivative values recorded among the species and showed less pronounced peaks across the spectrum, with minor fluctuations near 670 nm (Fig. 5B). Furthermore, at certain wavelengths, both positive and negative peaks observed for C1 were not present at other concentrations; for instance, at 475 nm and 650 nm, the C1 concentration showed a pronounced trough, C2 exhibited a slight trough, while the other concentrations seemed to trend in the opposite direction.

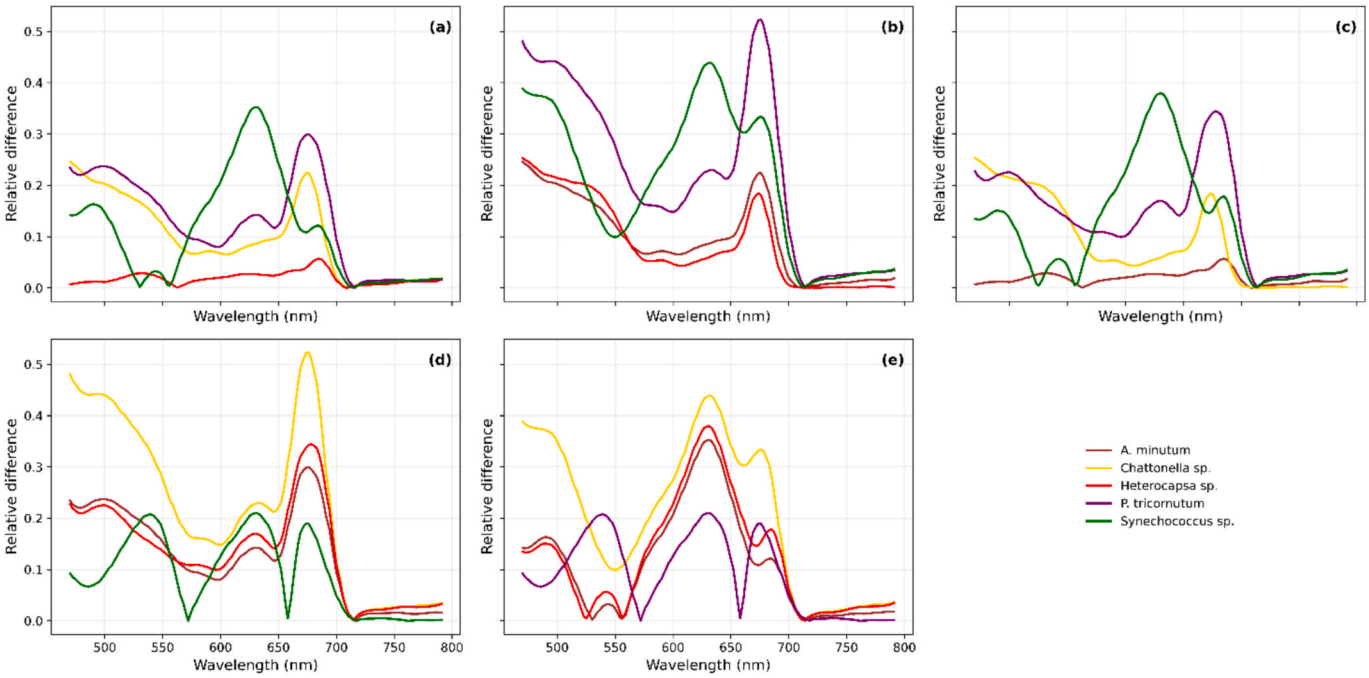


Fig. 4. Pairwise spectral differences between algal species. Each panel (a–e) shows the absolute spectral difference between one target species and all others, based on normalized and smoothed reflectance spectra. | The target species for each panel are as follows: (a) *Alexandrium minutum*, (b) *Chattonella* sp., (c) *Heterocapsa* sp., (d) *Phaeodactylum tricornutum*, and (e) *Synechococcus* sp.

P. tricornutum exhibited unique features between 600 and 640 nm, where a double peak is observed (Fig. 5D). In this spectral region, *Synechococcus* sp. showed a single large positive peak, while displaying negative values in the range between 520 and 560 nm, a region where the other species maintained positive values (Fig. 5E).

3.3. Similarity analysis

The Similarity Index (SI) analysis revealed how spectral profiles of different algae change with varying concentrations (Fig. 6).

Regarding *A. minutum*, this species generally showed the lowest SI values. In general, C2 showed the highest similarity to the reference spectrum (C1), and as the concentration decreased, the similarity index dropped significantly. For instance, in *A. minutum*, C2 showed a high similarity (SI of 0.55), C3 exhibited a moderate similarity with a SI of 0.41, while the lowest concentrations (C4, C5, and C6) displayed substantially lower similarity indices, all falling below 0.20, indicating pronounced spectral divergence at the lowest concentrations. *Chattonella* sp. followed a different pattern, with the highest similarity index at C2 (0.57) and a steady decline toward C5 (0.36). However, the decline was less pronounced than in the other microalgae, indicating that this species retained more consistent spectral characteristics even at intermediate concentrations. *P. tricornutum* exhibited a marked decrease in SI as concentrations decreased. The similarity dropped sharply from C2 (0.60) to C4 (0.14) and reached a minimum at C6 (0.07). Finally, *Synechococcus* sp. also showed a strong decline in similarity with decreasing concentrations, starting from C2 (0.60) and reaching its lowest at C6 (0.18). Here, the consistency between C1 and C2 is notable, but beyond C3, the similarity decreases substantially. In summary, across all species, the trend indicated that high concentrations (C2 and C3) generally yield higher similarity indices, reflecting more stable spectral features. However, as the concentration decreased, the spectral similarity diminished, particularly in species like *P. tricornutum, Heterocapsa* sp., and *Synechococcus* sp.

3.4. Dimensionality reduction and feature extraction

An initial PCA on the raw standardized reflectance data explained 93% of the total variance (Fig. 7A). However, the samples exhibited a pronounced linear alignment along the principal components, indicating that changes in overall reflectance intensity, likely due to cellular concentration and biovolume, dominated the observed variability. Loading analyses supported this conclusion, showing that PC1 primarily captured global intensity variations, while PC2 partially reflected the characteristic shape of the reflectance spectrum (Fig. 7B). To mitigate the influence of intensity-driven effects and emphasize spectral differences, a second PCA was conducted. Here, each sample's reflectance was normalized by its maximum value and an additional variable (species biovolume) was incorporated. This approach reduced the effect of overall intensity and allowed the PCA to focus more on the spectral differences among the species. The refined analysis explained 97% of the total variance (Fig. 7C), confirming that normalization decreased intensity-related effect and better highlighted the spectral features. Fig. 7D shows the loadings of the second PCA, where PC1 is no longer flat but displays more pronounced spectral patterns, confirming that the principal source of variability now reflects spectral differences among species, rather than intensity. This second PCA provided clearer discrimination among the algal species.

We then applied LDA and the results showed a clear separation of the five algal species in the reduced two-dimensional space, with distinct clusters for each species (Fig. 8).

The LDA accounted for 99.8% of the variance, with LDA Component 1 capturing 80.2% of the discriminatory power, suggesting that most of the inter-species' separation can be achieved by focusing on the first LDA component. Notably, *Synechococcus* sp. was clearly separated from the other taxa, reflecting its distinct pigment composition as a cyanobacterium. The LDA coefficients, initially computed on PCA-transformed spectrum, were projected back into the original wavelength space to estimate each band's contribution to species separation. This approach allowed us to rank the spectral bands by importance and identify the wavelengths most relevant for species discrimination (see Table S3). These wavelengths were predominantly located between 620 nm and

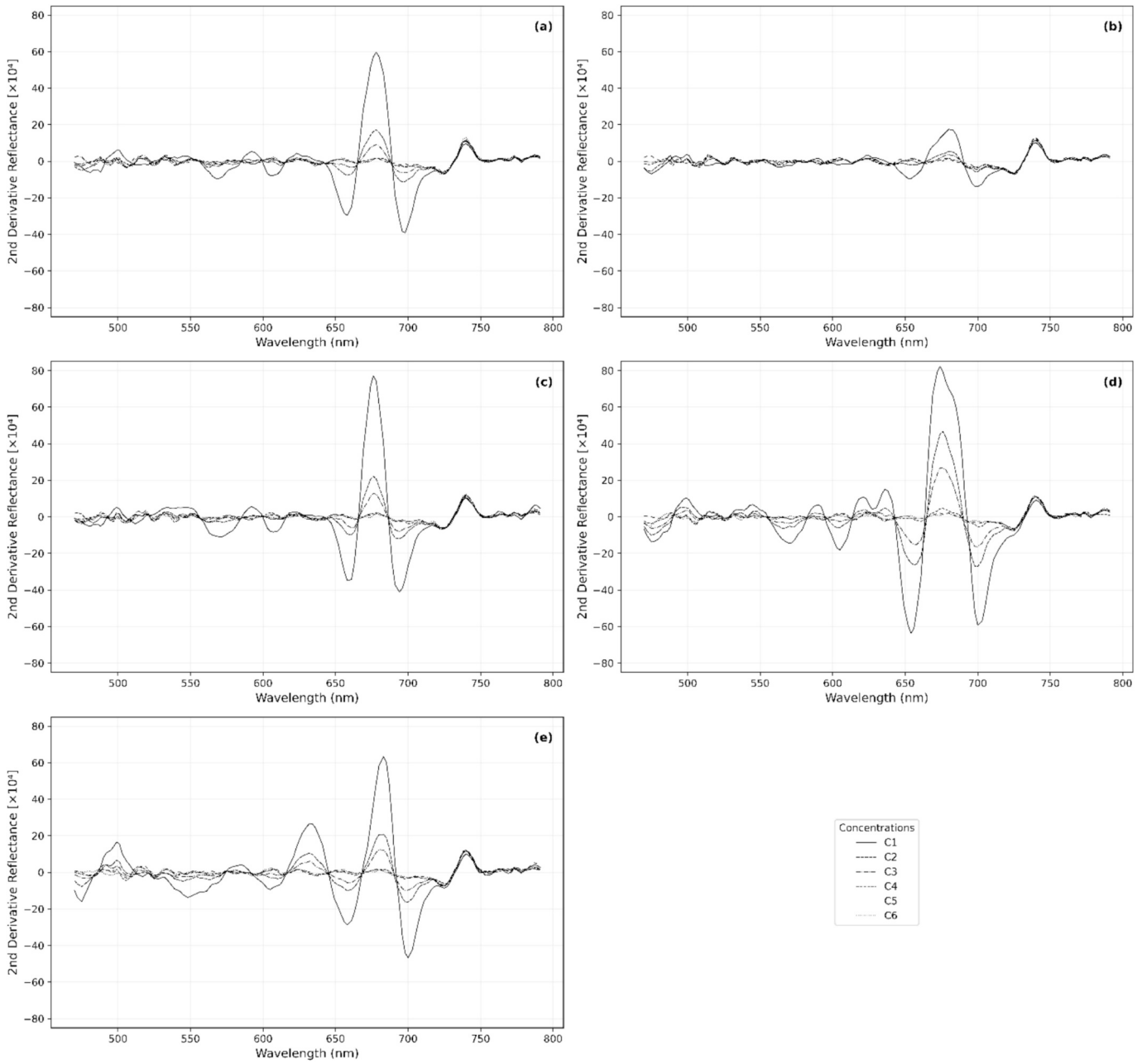


Fig. 5. Second derivative reflectance spectra across different cell concentrations. | Panels show: (a) *Alexandrium minutum*, (b) *Chattonella* sp., (c) *Heterocapsa* sp., (d) *Phaeodactylum tricornutum*, and (e) *Synechococcus* sp.

680 nm, a range known to be associated with key absorption features of algal pigments, such as chlorophylls and phycobiliproteins.

3.5. Spectral unmixing

Linear Spectral Unmixing has been applied to the hyperspectral reflectance data and provided estimates of fractional abundances for each algal species within three different algal mixtures (Mix 1, Mix 2, Mix 3). Estimated LSU fractions were compared against actual fractional abundances measured experimentally from biovolume proportions (Table 4). LSU performance varied across mixtures, with better agreement in simpler mixtures dominated by a single species and larger discrepancies in more complex mixtures; this performance degradation is most clearly captured by MAAE. Mix 1 (dominated by *Alexandrium minutum*) showed the best agreement between LSU-estimated fractions and measured biovolume (0.976 vs. 0.956), with a low RMSE (0.077) and a moderate SAM (5.1°), indicating a good spectral match with minor reconstruction errors.

For Mix 2, LSU estimated the fractional abundances of *P. tricornutum* to be nearly 100%, with negligible contributions from the other two species. This result aligned closely with the true biovolume-based composition of the Mix, where *P. tricornutum* accounted for approximately 93% of the total biovolume. Although the LSU model successfully identified the dominant species, it exhibited a RMSE of 0.376 and a SAM of 5.94°, indicating a moderate spectral mismatch between the reconstructed and the observed spectra. Mix 3, containing multiple co-dominant species (*A. minutum*, *Chattonella* sp., *P. tricornutum*, *Synechococcus* sp.), showed the largest deviations from measured biovolume, particularly for *A. minutum* (0.758 vs. 0.367). Despite this, RMSE (0.041) and SAM (3.82°) were low, indicating a good spectral fit even though abundance estimates were inaccurate. To quantitatively assess how LSU performance degraded with increasing mixture complexity, we focused

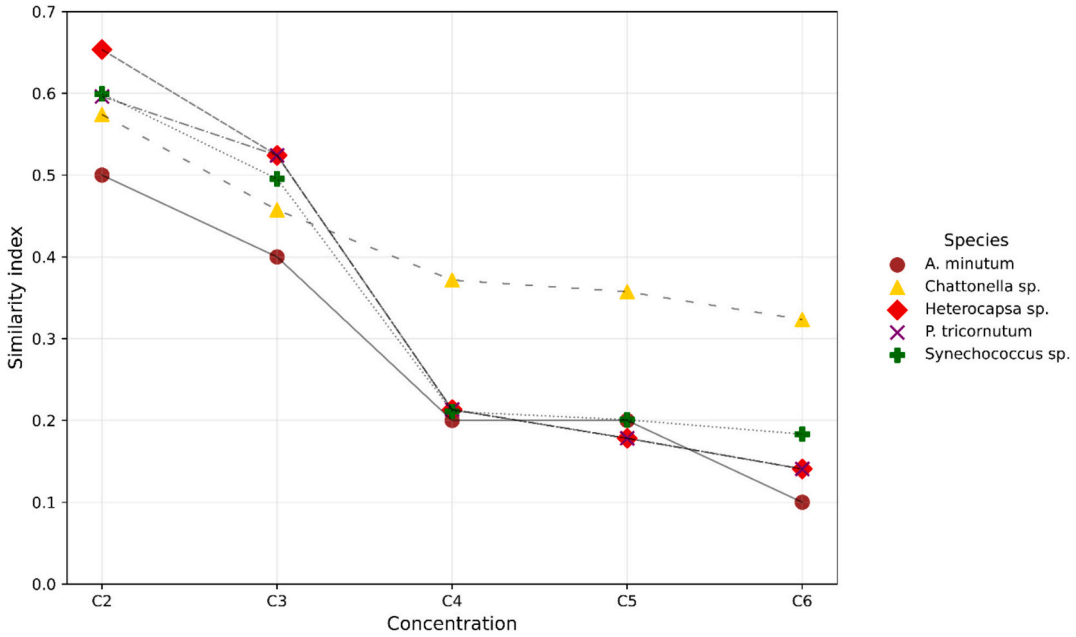


Fig. 6. Similarity Index trends across concentrations (C2-C6) for each algal species.

on the MAAE, as this metric directly captures inaccuracies in fractional abundance retrieval. In contrast, RMSE and SAM describe spectral reconstruction (magnitude and shape, respectively), which can remain low even when species abundances are poorly estimated, as observed in Mix 3. MAAE showed a clear and strong degradation pattern across mixtures. It increased from 0.020 in Mix 1 to 0.046 in Mix 2 (a 2.3-fold increase), and reached 0.256 in Mix 3, corresponding to a more than 12-fold increase relative to Mix 1.

4. Discussion

4.1. Spectral characteristics of phytoplankton species

We investigated the spectral signatures of five algal species at different cellular concentrations and observed that one of the main findings across all species was the presence of shifting peaks. These shifts could be due to low optical density and they might result from changes in the optical path length, enabling precise estimation of chlorophyll *a* concentration in diverse aquatic environments (GITELSON, 1992; Gitelson, 1993; Xi et al., 2017). The peaks shifted toward shorter wavelengths as concentration decreased. This is because at higher cellular concentrations light travels along shorter, more complex paths due to dense cell packing, which can cause shifts in peak positions. As the density lowers, light has a more straightforward path and may interact differently with remaining cells, shifting peak positions slightly toward shorter wavelengths. This shift can be especially noticeable in spectrum regions where pigments like chlorophyll exhibit strong fluorescence (Duan et al., 2009), such as in the "red edge" region between the red and near-infrared wavelengths (≈710 nm in our study), where reflectance undergoes a rapid change because chlorophyll fluorescence emission contributes to the reflected light, causing a slight increase in apparent reflectance (Zarco-Tejada, 2000). Regarding the general analysis of the individual spectra of the five phytoplankton species, reflectance revealed distinct optical behaviors; Table 5 provides a summary list of the specific wavelengths at which reflectance peaks and absorption features were observed in the algae tested in this study.

Although some species exhibited partially overlapping absorption and reflectance ranges, particularly within pigment-dominated regions such as 470–500 nm and around 680 nm, these similarities mainly arise from the shared presence of chlorophyll *a* and accessory carotenoids. Consequently, species with comparable pigment compositions (e.g., *A. minutum* and *Heterocapsa* sp.) may show similar absorption features, potentially leading to partial spectral interference when relying solely on reflectance intensity. However, differences in the overall spectral shape, including band depth, width, slope, and spectral derivative still allow discrimination when the full hyperspectral signature is considered. Multivariate analyses such as PCA and LDA, as applied in this study, further enhance class separability by accounting for subtle variance patterns across the entire spectrum.

The spectral overlap issue is well recognized in hyperspectral phytoplankton studies, as microalgal optical properties are influenced by cell morphology, size, and pigment organization, all of which can introduce nonlinear light-scattering effects (Bricaud et al., 2007; Mehrubeoglu et al., 2014; Salmi et al., 2022). While physics-based modeling or non-linear unmixing could improve accuracy, these approaches require extensive optical characterization of each species. Alternatively, recent work has highlighted the potential of machine-learning models to resolve complex, non-linear relationships in hyperspectral data (Salmi et al., 2022; Solovchenko, 2023), but such techniques demand large, standardized reference datasets for reliable training, beyond the scope of controlled laboratory studies like the present one. Our results provide empirical evidence that dense VNIR hyperspectral sampling is critical for resolving phytoplankton taxa with partially overlapping pigment absorption features. Although the two dinoflagellates (*Alexandrium minutum* and *Heterocapsa* sp.) share similar chlorophyll- and carotenoid-driven absorption regions (e.g., 470–500 nm and near 680 nm), we achieved consistent separation by leveraging subtle differences in spectral shape (band depth/width and slopes) across the full spectrum, together with derivative-based and multivariate analyses. This capability is expected to be reduced for multispectral sensors that sample only a limited number of broad, discontinuous bands, where such subtle features may be smoothed or missed (Braga et al., 2022; Dierssen et al., 2021; Jaywant and Arif, 2024)

Each species exhibited unique spectral signatures that are reflective of differences in their cellular structure, pigment composition, and light interaction strategies. We investigated two complementary normalization approaches to understand the optical properties of the species and we applied the second derivative to the reflectance spectra. This enhanced the detection of minor absorption features that were not apparent in the original spectra (see supplementary materials, Fig. S3,

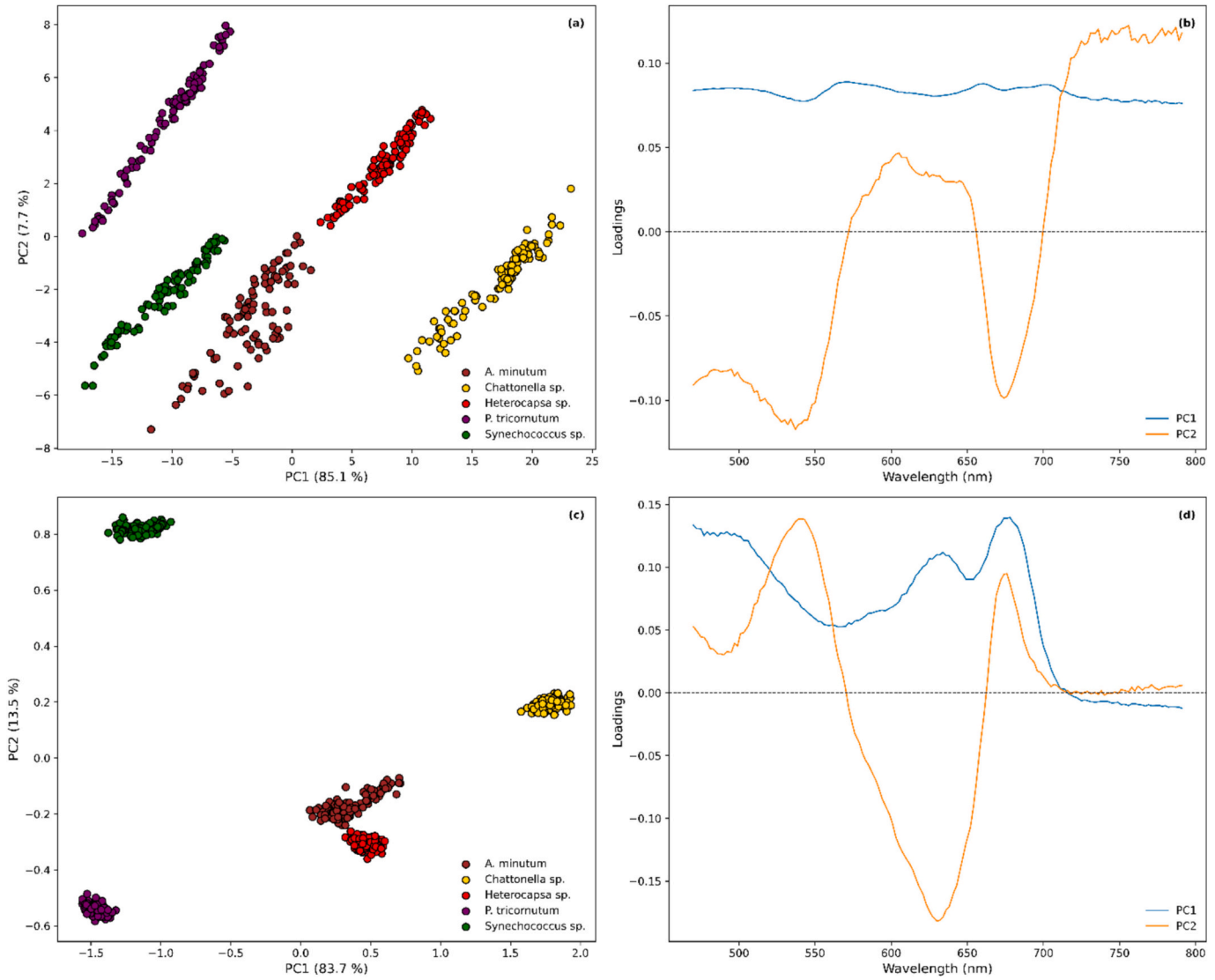


Fig. 7. Principal-component analysis of reflectance spectra. | (a) Scores of the first PCA computed on z-standardized raw spectra. (b) Corresponding loadings for PC1 and PC2. (c) Scores of a second PCA after within-species maximum-normalization and inclusion of biovolume. (d) Loadings for PC1 and PC2 of the second PCA, indicating the wavelength regions that contribute most to species separation.

for more information about second derivative).

The second derivative was particularly effective in isolating the spectral signatures of specific pigments and it has been used in different papers to retrieve phytoplankton optical signatures and pigments (Aguirre-Gómez et al., 2001; Lavigne et al., 2022; Ruddick et al., 2023). Additionally, it is known to help reduce the influence of light scattering and noise caused by cell morphology and internal structure, thus allowing for a clearer identification of pigment-related absorption peaks (Ruddick et al., 2023). The first normalization approach was based on biovolume and emphasized the efficiency of light interaction per unit biomass. *Chattonella* sp. and *Heterocapsa* sp. exhibited higher reflectance per biovolume, suggesting they could scatter light more efficiently; this is particularly notable for *Chattonella* sp., as these organisms lack a cell wall and can assume various cellular shapes depending on the environmental conditions they encounter (Klöpper et al., 2013). In contrast, *A. minutum* had lower reflectance per biovolume, implying either denser cells or more efficient light absorption due to high pigment concentrations. Laboratory studies have shown that *A. minutum* exhibits high chlorophyll *a* content and elevated photosynthetic efficiency, particularly under low light intensities (Li et al., 2021). These findings suggest that the lower reflectance observed in our study could be attributed to an enhanced capacity for light absorption, possibly facilitated by its high pigment concentration and efficient use of available light. This species displayed a relatively flat spectral profile with lower reflectance efficiency, which may indicate either denser cellular structures or a broader absorption spectrum without selective peaks. The same observations could be made for *P. tricornutum* and *Synechococcus* sp., which both display very low reflectance per biovolumes. The other normalization approach was based on scaling each spectrum to its maximum value to emphasize the relative spectral shape without taking cellular concentration or biomass into account, but focusing on absorption features only. Notably, the absorption peak around 680 nm was present across species; that position corresponds to the absorption range of chlorophyll *a*, the dominant pigment common to all microalgae. Another common trend observed across species (Table 5) is the strong absorption in the blue-to-green region, approximately between 470 and 500 nm (see also Fig. S3). This area is characterized by multiple overlapping absorption bands due to the presence of various pigments, as all carotenoids are known to absorb most efficiently in this region (Luimstra et al., 2020; Roy et al., 2011). Moreover, chlorophyll *a* also has a secondary absorption peak in this wavelength range, which may contribute to the overall high absorption. However, given the vast diversity of pigments,

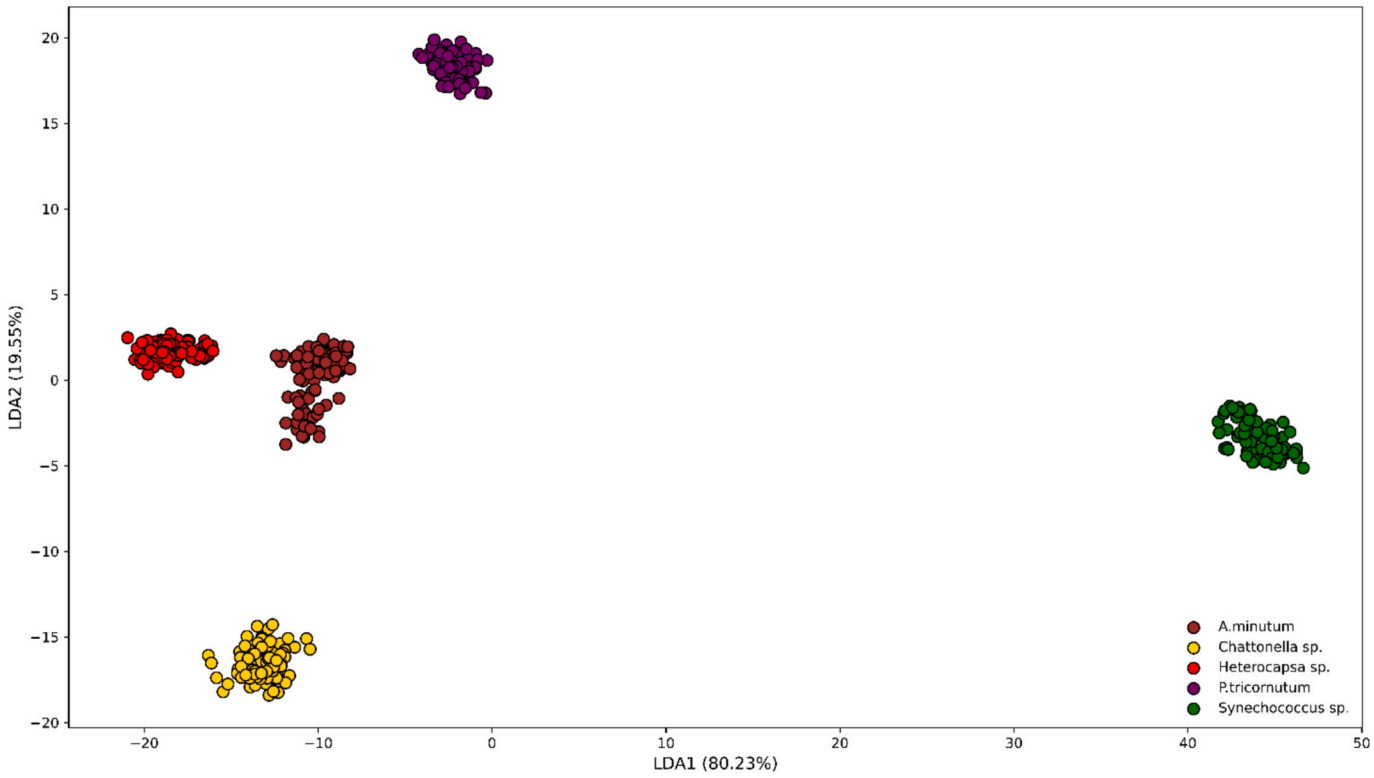


Fig. 8. Linear Discriminant Analysis (LDA) applied to the first three principal components of the standardized reflectance spectra from five phytoplankton species.

Table 4
Comparison between LSU-estimated fractional abundances and measured biovolume proportions of algal species in the three different mixtures, along with RMSE, SAM and MAAE values assessing the accuracy of the LSU reconstruction.

Mixture	Algal Species	LSU-estimated fraction	Actual Fraction (measured biovolume)	RMSE	SAM (degrees)	MAAE
Mix 1	*Alexandrium minutum*	0.976	0.956	0.077	5.1	0.020
	Heterocapsa sp.	0.024	0.044			
Mix 2	*Heterocapsa* sp.	≈ 0.000	0.058	0.376	5.94	0.046
	Phaeodactylum tricornutum	≈ 1.000	0.931			
	Synechococcus sp.	≈ 0.000	0.011			
Mix 3	*Alexandrium minutum*	0.758	0.367	0.041	3.82	0.256
	Chattonella sp.	0.118	0.367			
	Phaeodactylum tricornutum	≈ 0.000	0.263			
	Synechococcus sp.	0.123	0.003			

Table 5
List of absorption and reflectance ranges for the tested species.

Phylum	Species	Harmful impact	Absorbance range	Reflectance range
Dinoflagellata	*Alexandrium minutum*	HABs, PSP toxin	470–500 583–594 627–643 658–691	572–582 596–616 641–645 696–722
Ochrophyta	*Chattonella* sp.	HABs, Ichthyotoxic	470–500 616–632 658–691	594–607 634–654 694–718
Dinoflagellata	*Heterocapsa* sp.	Potentially harmful	470–500 661–687	599–618 694–722
Bacillariophyta	*Phaeodactylum tricornutum*	Non-toxic, bloom-forming	470–510 627–641 658–694	592–612 641–654 698–727
Cyanobacteria	*Synechococcus* sp.	Non-toxic, bloom-forming	470–500 621–643 669–683	545–565 652–665 696–727

this spectral region presents challenges for distinguishing specific species. The absorbance contributed by accessory photopigments is often difficult to quantify and reliably differentiates from the 'masking effect' of chlorophyll *a*, as well as the inherent 'background' absorbance resulting from cell packaging effects (Millie et al., 2002). On the other hand, at around 710 nm, all species exhibit a strong reflectance peak. This could be attributable to chlorophyll *a*, which may further contribute to the increase in reflectance in this zone due to secondary effects such as re-emission of absorbed light or fluorescence (Chen et al., 2021). Cell scattering could also influence this part of the spectrum, especially because complex cellular structures can significantly affect light scattering in regions where pigments do not absorb light. To accurately assess these effects, it is essential to expand the analysis to include quantification of pigments and measurement of cellular scattering. In our study, *A. minutum* and *Heterocapsa* sp. exhibited a similar pattern, indeed, these species belong to the same group of microalgae, both dinoflagellates. The pattern observed in both algae shows a strong absorption zone in the blue region, likely due to the presence of characteristic carotenoids such as peridinin, diadinoxanthin, and dinoxanthin, along with a mix of chlorophylls (*a* and *c2*) which strongly absorb light in this region (Table 1). The green-to-yellow region

(500–600 nm) displays higher reflectance, generally corresponding to a lower presence of absorbing pigments. In the reflectance curves of both, a small reflectance peak is noticeable between 600 and 615 nm. This can be explained by the lack of absorption molecules in the area. In fact, chlorophyll *c2* has a typical absorption peak between 580 and 590 nm and the gap between this absorption and that of chlorophyll *a* leads to increased reflectance in the spectra. In contrast, *Chattonella* sp. exhibited slightly different behaviors; the absorption in the blue region remains prominent and is primarily attributed to chlorophyll *a* and fucoxanthin, which are commonly present in this species. The small peak observed at 650 nm could be due to scattering, which might be higher in this region due to cell structure and the lack of absorbing pigments. The gradual increase in reflectance from blue to yellow may indicate that this species uses a combination of carotenoids and cellular adaptations to optimize light absorption without creating distinct spectral signatures. Alternatively, it could suggest that our cultured species lacked a high concentration of pigments. Conversely, *P. tricornutum* showed pronounced peaks and troughs, likely due to the presence of accessory pigments such as fucoxanthin, which is the dominant pigment in diatoms (Roy et al., 2011) and chlorophyll *a*. The increase in the reflectance between 600 and 615 nm can be attributed to the same reasons observed in other species. However, a marked trough at 630 nm is also present; this feature is probably linked to the presence of chlorophyll *c*, which has a secondary absorption window in this area. In the environment this pigment plays a crucial role in expanding the light absorption capabilities of diatoms, allowing them to thrive under various light conditions (Zhou et al., 2021). Regarding *Synechococcus* sp., a cyanobacterium, this species exhibits distinctive and characteristic spectral features (Table 5). The absorption between 470 and 500 nm corresponds to typical absorption features also observed in other algae, arising from the presence of various pigments that absorb light at these wavelengths. Furthermore, this region also exhibits strong absorption due to the presence of phycobiliproteins, which are accessory pigments specific to cyanobacteria. The rapid increase in reflectance up to 550 nm may indicate the absence of absorbing pigments in this region. Notably, there is strong absorption between 620 and 630 nm, primarily due to the presence of phycocyanin, which is usually highly abundant in cyanobacteria. As previously noted, cyanobacteria exhibit significant optical differences from other phytoplankton groups and are therefore more easily identifiable using hyperspectral data (Xi et al., 2015). This absorption band allows *Synechococcus* sp. to utilize wavelengths that are less exploited by other phytoplankton, enhancing its ability to thrive in low-light environments or deeper water layers (Kolodny et al., 2022). Having observed these unique spectral characteristics for each species, we wanted to analyze whether and how these spectral features remain distinct as cell concentration decreases by applying a SI. Overall, at lower concentrations, the spectral signature deviated from the reference spectra, likely due to a reduction in absorption features and an increase in background noise dominated by scattering effects. As observed in our analysis, *P. tricornutum* exhibited a particularly strong decline in spectral consistency with decreasing cell concentration, as its spectral properties are highly sensitive to changes in concentration, especially at lower cell densities. These results are in line with those presented by Soja-Woźniak et al. (2018), where the SI values were found to be very similar to ours when the ratios between the reference concentrations and other concentrations were close to those tested in our study. This suggests that high cell densities are necessary to obtain reliable reflectance measurements, as lower concentrations can obscure species-specific spectral features, complicating the differentiation of mixed samples (Moloto et al., 2023). In natural aquatic environments such as the Adriatic Sea, typical phytoplankton concentrations often fall below the levels used in this study settings, especially under non-bloom conditions, making spectral discrimination more challenging. However, during algal bloom events, cell densities increase significantly, potentially increasing the accuracy of remote sensing identification.

4.2. PCA and LDA processes

The use of PCA and LDA provided complementary insights into the dataset. The initial PCA and the loadings revealed a strong linear alignment among the samples, indicating that overall intensity differences driven by variations in concentration and biovolume were the primary sources of variance. In hyperspectral data, the high collinearity between wavelength bands could highlight a substantial spectral redundancy (Cai et al., 2023; Fu et al., 2024; Kumar et al., 2020). This could end up overloading the analysis, requiring more processing time, and potentially causing confusion. In literature, several dimensionality reduction methods have been explored, and among them, PCA has proven to be an excellent technique to use with hyperspectral data (Vaddi et al., 2024). The wavelengths with the highest discriminatory power were identified by projecting the LDA coefficients back into the original spectral space. This highlighted the most discriminative wavelengths between 620 and 680 nm region and at 550 nm and 600 nm, corresponding to key pigments like chlorophyll *a*, phycocyanin, and fucoxanthin (see supplementary, Table S2). The prominence of these bands suggests their potential utility in automated classification systems for algal species based on spectral data. Interestingly, the results of LDA and the second PCA were highly similar, suggesting that after accounting for biovolume normalization, PCA alone was sufficient to separate the species based on their spectral profiles. While the refined PCA improved species separation, it remains an unsupervised method focused on variance, not necessarily on optimizing class separation. The LDA, in this context, served more as a validation step, confirming that the separation observed in the PCA was indeed robust and biologically meaningful. The integration of spectral analysis, normalization techniques, PCA, and LDA provided a multi-faceted view of phytoplankton optical properties. The results demonstrated that while intensity-driven differences can dominate initial spectral analyses, proper normalization and dimensionality reduction techniques can effectively isolate species-specific features. These methodologies not only improve our understanding of phytoplankton ecology but also enhance the ability to utilize remote sensing for environmental monitoring.

4.3. Unmixing approach

The LSU method is widely used in remote sensing and hyperspectral analysis due to its simplicity and interpretability (Liu et al., 2016). Although nonlinear unmixing techniques and machine learning-based approaches have been proposed, they were not selected for this study for several key reasons. In our controlled algal suspensions, only reflectance data were available, and accurately modeling nonlinear interactions, such as multiple scattering or fluorescence re-emission, would require additional calibration and absorption–scattering data (Cavalli, 2022).

Given the experimental constraints and the need for transparent, physically meaningful results, a linear approach was then considered the most appropriate. Machine learning models, while powerful, generally require large training datasets and often act as “black boxes,” making it difficult to interpret the physical meaning of the estimated abundances (Allen et al., 2025; Azodi et al., 2020). In this study, the dataset was limited to a certain number of pixels per species from homogeneous samples, making machine learning approaches impractical. In this context, LSU provided reliable abundance estimates in the simpler mixtures dominated by one or two species (Mix 1 and Mix 2), confirming its robustness for low-complexity assemblages. For example, in Mix 2, *Phaeodactylum tricornutum* was correctly identified as the dominant species, accounting for nearly all of the modeled reflectance signal. However, the relatively high RMSE (0.376) suggests that the strong spectral signature of *P. tricornutum* may have masked weaker contributions from the minor components (*Heterocapsa* sp., *Synechococcus* sp.). This is a known limitation of linear mixture models that tend to underestimate minor endmembers when their inclusion does not

substantially improve the spectral fit. Despite this, the low SAM value (5.9°) indicates that the reconstructed spectrum closely matched the observed spectral shape, confirming that LSU successfully reproduced the main spectral features even if abundance estimates were biased.

In contrast, Mix 3, containing several co-dominant species with overlapping spectral features, highlighted the limits of LSU in more complex assemblages. Here, *A. minutum* was strongly overestimated (0.758 vs. 0.367 actual fraction), whereas the overall spectral reconstruction remained accurate (RMSE = 0.041; SAM = 3.8°). This apparent paradox reflects LSU's ability to find mathematically valid combinations that fit the measured spectrum while misrepresenting the underlying biological composition, emphasizing that LSU accuracy decreases as spectral overlap among species increases, particularly when several taxa contribute similarly to the total reflectance signal.

Consistent with this pattern, the MAAE revealed a substantial escalation in abundance estimation error, demonstrating that LSU increasingly failed to recover the correct abundance structure as mixtures became more balanced and spectrally heterogeneous. Specifically, MAAE increased more than twelve-fold from Mix 1 to Mix 3, underscoring the rapid loss of abundance-retrieval accuracy as mixture complexity and spectral overlap increase. Unlike RMSE and SAM, which may indicate acceptable spectral fits even when abundances are incorrect, MAAE provides a more reliable and ecologically meaningful measure of LSU performance in complex algal compositions.

Although phytoplankton communities in natural waters can vary widely and often include multispecific assemblages with several abundant species (Guinder et al., 2015; Shen et al., 2019; Zhang et al., 2016), our experimental design was not intended to reproduce the full complexity of such systems. Instead, it was developed to reflect ecological scenarios relevant for harmful algal blooms (HABs) detection, where one or a few taxa typically reach high dominance. At the same time, we acknowledge that in many productive coastal environments, multiple species may also coexist at relatively high densities. In these situations, increased spectral overlap among co-occurring taxa can reduce spectral separability, making abundance estimation more challenging for linear unmixing approaches.

Real-world spectra are further complicated by other optically active constituents such as colored dissolved organic matter (CDOM), suspended sediments and detritus, which can mask or distort algal signatures and may require additional correction before LSU (or any unmixing approach) is applied. This highlights a potential bias in LSU toward species with strong spectral contrasts when multiple similar species are present.

Additionally, intrinsic differences in reflectance intensities, pigment composition, and cellular structures can contribute to spectral interference in complex mixtures. While LSU provides a useful tool for reasoning about the estimation of algal abundances, its accuracy ultimately depends on community composition, species dominance and spectral interactions. Future research should focus on extending datasets and considering advanced unmixing or hybrid models capable of capturing nonlinearities and the true variability of such complex interactions. Refining these approaches will enhance the reliability of hyperspectral unmixing for monitoring complex algal communities, as well as bloom-forming events, in natural aquatic environments.

4.4. Potential and limitations

Hyperspectral reflectance analysis has demonstrated its potential in distinguishing species based on their unique optical signatures (McKibben et al., 2024; Zhang et al., 2024), particularly in controlled laboratory settings, where factors such as light conditions, cell concentration, and spectral noise can be carefully managed.

The ability to analyze spectral features at high resolution makes hyperspectral imaging a valuable tool for monitoring algal blooms, tracking phytoplankton community shifts, and assessing ecosystem health across large spatial scales (Solovchenko, 2023). Additionally, advances in hyperspectral imaging have improved species discrimination based on pigment composition and cellular responses, crucial for early HABs detection and environmental monitoring (Dierssen et al., 2021). Despite these advantages, there are significant limitations to using reflectance alone for phytoplankton monitoring, particularly in natural environments. Reflectance spectra are influenced not only by pigment absorption but also by light scattering, which is affected by cell size, shape and internal structure (Vaillancourt, 2004; Whitmire et al., 2010), and without prior knowledge of these factors, interpreting reflectance data can be challenging. The presence of other optically active components, such as CDOM and non-algal particles/suspended sediments, can substantially affect the magnitude and shape of the observed reflectance (including water-leaving reflectance in the field) and progressively mask phytoplankton-driven spectral features, particularly at shorter wavelengths (<550 nm), thereby complicating spectral interpretation in natural waters (Zhang et al., 2024).

Moreover, our SI analysis indicates that spectral consistency decreases at lower cell concentrations, suggesting that sub-bloom phytoplankton densities typical of many coastal conditions may weaken diagnostic absorption features and increase the relative contribution of background scattering and noise. In addition, our unmixing results showed a marked performance degradation as mixtures became more complex and spectrally overlapping, highlighting that abundance retrieval becomes increasingly uncertain when multiple taxa contribute comparably to the observed signal. Together, these factors imply that algorithms calibrated under controlled laboratory conditions require careful validation under realistic field variability. Therefore, extensive and well-characterized ground-truth datasets (spanning seasons, trophic conditions and community compositions) are crucial to quantify uncertainty, validate classification/unmixing performance and support robust generalization to operational monitoring.

Additionally, external factors like cloud cover, water depth, and surface reflection introduce variability in reflectance, affecting chlorophyll *a* estimation accuracy and species differentiation (Kutser, 2004). Moreover, phytoplankton organisms are part of complex ecological communities that include bacteria, zooplankton, and other photoautotrophic organisms. These interactions can influence their optical properties, as well as their growth and pigment composition.

As a result, future efforts should prioritize coupling laboratory-derived optical libraries with field measurements to refine model parameterization and enable operational monitoring tools. Machine learning and non-linear unmixing techniques could further enhance classification performance, provided that sufficient training data are available to constrain and validate model predictions (Salmi et al., 2022; Zhu et al., 2019).

To improve accuracy, it is essential to integrate reflectance measurements with additional optical parameters, such as absorption and scattering coefficients, as well as information on the cellular structure. Moreover, hyperspectral data should be collected at high temporal frequency, since phytoplankton communities can change rapidly in response to environmental fluctuations, nutrient availability and physical mixing processes (Henson et al., 2021; Sjöqvist, 2022). Frequent acquisitions are therefore crucial to accurately capture short-term variations and dynamic shifts in community composition (Dierssen et al., 2021; Lavigne et al., 2022). Furthermore, the application of machine learning algorithms to hyperspectral datasets shows promise in improving classification accuracy, even in the presence of mixed signals from natural waters (Zhu et al., 2019). By using large datasets and training models on both lab and field data, it may be possible to develop more robust algorithms that account for the variability in natural environments.

In this study, dense VNIR hyperspectral sampling proved particularly valuable for discriminating taxa with partially overlapping pigment absorption features, by preserving subtle differences in spectral shape and supporting derivative-based and multivariate analyses. However, the operational transfer of these approaches will depend on robust field

validation and on accounting for the confounding effects discussed above (Braga et al., 2022; Fabbretto et al., 2024; Jaywant and Arif, 2024).

5. Conclusion

This study investigated the spectral characteristics of different phytoplankton species belonging to classes commonly found in coastal waters, applying advanced dimensionality reduction techniques and spectral unmixing methods to enhance species differentiation.

Overall, hyperspectral analysis combined with normalization, derivative-based features and multivariate methods (PCA/LDA) enabled the identification of species-specific optical patterns relevant to environmental monitoring. However, closely related taxa (e.g., the two dinoflagellates) exhibited highly similar spectral behavior, underscoring the ongoing challenge of species-level discrimination using reflectance alone. Future developments should therefore focus on expanding hyperspectral reference datasets and coupling laboratory spectral libraries with in situ ground-truth measurements to validate and generalize classification and unmixing performance under realistic field variability. Ultimately, these advances could support the integration of hyperspectral analyses into predictive models and early-warning systems for HAB monitoring.

CRediT authorship contribution statement

R. Bentivogli: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **L. Pezzolesi:** Writing – review & editing, Supervision, Methodology, Conceptualization. **N. Caputo:** Investigation, Conceptualization. **B. Casarotto:** Writing – review & editing, Software, Investigation. **S. Silvestri:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Funding

This work was carried out within the RETURN Extended Partnership and received funding from the European Union – NextGenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005).

The study was also supported by the CASCADE project financed by Cooperation Program Interreg V-A Italia-Croazia 2014/2020.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the staff of the Department of Geosciences, University of Padova for technical assistance during the experimental setup and data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2026.103626.

Data availability

The datasets and scripts generated and used in this study are publicly available on Zenodo (DOI: 10.5281/zenodo.17563974). This DOI refers to the versioned snapshot of the materials used to produce the results in the manuscript.

References

Adejimi, O.E., Sadhasivam, G., Schmilovitch, Z., Shapiro, O.H., Herrmann, I., 2023. Applying hyperspectral transmittance for inter-genera classification of cyanobacterial and algal cultures. Algal Res. 71, 103067. https://doi.org/10.1016/j.algal.2023.103067.

Aguirre-Gómez, R., Weeks, A.R., Boxall, S.R., 2001. The identification of phytoplankton pigments from absorption spectra. Int. J. Remote Sens. 22, 315–338. https://doi.org/10.1080/014311601449952.

Allen, G.I., Gan, L., Zheng, L., 2025. Annual Review of Statistics and its Application Interpretable Machine Learning for Discovery: Statistical Challenges and Opportunities. On: Wed. https://doi.org/10.1146/annurev-statistics-040120.

Aparicio-Rizzo, P., Poblete-Caballero, D., Vera-Bastidas, C., Pérez-Santos, I., Varela, D., 2025. Phytoplankton detection study through hyperspectral signalling in the Patagonian fjords. Ocean Sci. 21, 2379–2395. https://doi.org/10.5194/os-21-2379-2025.

APAT, IRSA-CNR, 2003. Determinazione della clorofilla: metodo spettrofotometrico (Metodo 9020). In: Metodi analitici per le acque. Manuali e Linee Guida, 29/2003, 3, pp. 1137–1142.

Azodi, C.B., Tang, J., Shiu, S.H., 2020. Opening the black box: interpretable machine learning for geneticists. Trends Genet. https://doi.org/10.1016/j.tig.2020.03.005.

Bi, X., Lin, S., Zhu, S., Yin, H., Li, Z., Chen, Z., 2019. Species identification and survival competition analysis of microalgae via hyperspectral microscopic images. Optik (Stuttg). 176, 191–197. https://doi.org/10.1016/j.ijleo.2018.09.077.

Braga, F., Fabbretto, A., Vanhellemont, Q., Bresciani, M., Giardino, C., Scarpa, G.M., Manfè, G., Concha, J.A., Brando, V.E., 2022. Assessment of PRISMA water reflectance using autonomous hyperspectral radiometry. ISPRS J. Photogramm. Remote Sens. 192, 99–114. https://doi.org/10.1016/j.isprsjprs.2022.08.009.

Bresciani, M., 2011. Recognizing harmful algal bloom based on remote sensing reflectance band ratio. J. Appl. Remote. Sens. 5, 053556. https://doi.org/10.1117/1.3630218.

Bricaud, A., Claustre, H., Ras, J., Oubelkheir, K., 2004. Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. J. Geophys. Res. Oceans 109. https://doi.org/10.1029/2004JC002419.

Bricaud, A., Mejia, C., Blondeau-Patissier, D., Claustre, H., Crepon, M., Thiria, S., 2007. Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons. Appl. Opt. 46, 1251. https://doi.org/10.1364/AO.46.001251.

Cai, W., Ning, X., Zhou, G., Bai, X., Jiang, Y., Li, W., Qian, P., 2023. A novel hyperspectral image classification model using bole convolution with three-direction attention mechanism: small sample and unbalanced learning. IEEE Trans. Geosci. Remote Sens. 61, 1–17. https://doi.org/10.1109/TGRS.2022.3201056.

Čanković, M., Dutour-Sikirić, M., Radić, I.D., Ciglenečki, I., 2022. Bacterioneuston and bacterioplankton structure and abundance in two Trophically distinct marine environments — a marine Lake and the adjacent coastal site on the Adriatic Sea. Microb. Ecol. 84, 996–1010. https://doi.org/10.1007/s00248-021-01934-1.

Cavalli, R.M., 2022. Spatial validation of spectral unmixing results: a case study of Venice City. Remote Sens. 14. https://doi.org/10.3390/rs14205165.

Chen, H., Li, K., Xue, C., Wang, Q., 2021. A novel method for non-invasive estimation of primary productivity in aquatic ecosystems using a chlorophyll fluorescence-induced dynamic curve. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.682250.

Cuartero, A., Cáceres-Merino, J., Torrecilla-Pinero, J.A., 2023. An application of C2-net atmospheric corrections for chlorophyll-a estimation in small reservoirs. Remote Sens. Appl. 32. https://doi.org/10.1016/j.rsase.2023.101021.

Dashkova, V., Malashenkov, D.V., Baishulakova, A., Davidson, T.A., Vorobjev, I.A., Jeppesen, E., Barteneva, N.S., 2022. Changes in phytoplankton community composition and phytoplankton cell size in response to nitrogen availability depend on temperature. Microorganisms 10. https://doi.org/10.3390/microorganisms10071322.

Davidson, K., Gowen, R.J., Harrison, P.J., Fleming, L.E., Hoagland, P., Moschonas, G., 2014. Anthropogenic nutrients and harmful algae in coastal waters. J. Environ. Manag. 146, 206–216. https://doi.org/10.1016/j.jenvman.2014.07.002.

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. Front. Environ. Sci. https://doi.org/10.3389/fenvs.2021.649528.

Ding, W., Li, C., 2024. Algal blooms forecasting with hybrid deep learning models from satellite data in the Zhoushan fishery. Ecol. Inform. 82. https://doi.org/10.1016/j.ecoinf.2024.102664.

Dorantes-Aranda, J.J., 2023. Harmful algae impacting aquatic organisms: recent field and laboratory observations. Toxins (Basel). https://doi.org/10.3390/toxins15050339.

Duan, H.-T., Ma, R.-H., Zhang, Y.-Z., Zhang, B., 2009. Fluorescence peak shift corresponding to high chlorophyll concentrations in inland water. Guang Pu Xue Yu Guang Pu Fen Xi 29, 161–164.

Fabbretto, A., Bresciani, M., Pellegrino, A., Alikas, K., Pinardi, M., Mangano, S., Padula, R., Giardino, C., 2024. Tracking water quality and macrophyte changes in Lake Trasimeno (Italy) from spaceborne hyperspectral imagery. Remote Sens. 16. https://doi.org/10.3390/rs16101704.

Feng, L., Hu, C., Han, X., Chen, X., Qi, L., 2015. Long-term distribution patterns of chlorophyll-a concentration in China's largest freshwater lake: MERIS full-resolution observations with a practical approach. Remote Sens. 7, 275–299. https://doi.org/10.3390/rs70100275.

Fournier, C., Quesada, A., Cirés, S., Saberioon, M., 2024. Discriminating bloom-forming cyanobacteria using lab-based hyperspectral imagery and machine learning:

validation with toxic species under environmental ranges. Sci. Total Environ. 932. https://doi.org/10.1016/j.scitotenv.2024.172741.
Fu, B., Sun, X., Cui, C., Zhang, J., Shang, X., 2024. Structure-preserved and weakly redundant band selection for hyperspectral imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 17, 12490–12504. https://doi.org/10.1109/JSTARS.2024.3425906.
Gitelson, A., 1992. The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. Int. J. Remote Sens. 13, 3367–3373. https://doi.org/10.1080/01431169208904125.
Gitelson, A.A., 1993. In: Oron, M., Shladov, I., Weissman, Y. (Eds.), Nature of the Peak near 700 nm on the Radiance Spectra and its Application for Remote Estimation of Phytoplankton Pigments in Inland Waters, p. 170. https://doi.org/10.1117/12.150992.
Gower, J., King, S., Goncalves, P., 2008. Global monitoring of plankton blooms using MERIS MCI. In: International Journal of Remote Sensing. Taylor and Francis Ltd, pp. 6209–6216. https://doi.org/10.1080/01431160802178110.
Grunert, B., Ciochetto, A., Mouw, C., 2025. A hyperspectral approach for retrieving inherent optical properties, phytoplankton pigments, and associated uncertainties from non-water absorption. Front. Mar. Sci. 12. https://doi.org/10.3389/fmars.2025.1549312.
Guillard, R.R.L., Ryther, J.H., 1962. Studies of marine planktonic diatoms: I. Cyclotella Nana Hustedt, and Detonula Confervacea (CLEVE) gran. Can. J. Microbiol. 8, 229–239. https://doi.org/10.1139/m62-029.
Guinder, V.A., López-Abbate, M.C., Berasategui, A.A., Negrin, V.L., Zapperi, G., Pratolongo, P.D., Fernández Severini, M.D., Popovich, C.A., 2015. Influence of the winter phytoplankton bloom on the settled material in a temperate shallow estuary. Oceanologia 57, 50–60. https://doi.org/10.1016/j.oceano.2014.10.002.
He, X., Liu, Y., Ganesan, K., Ahnood, A., Beckett, P., Eftekhari, F., Smith, D., Uddin, M.H., Skafidas, E., Nirmalathas, A., Unnithan, R.R., 2020. A single sensor based multispectral imaging camera using a narrow spectral band color mosaic integrated on the monochrome CMOS image sensor. APL Photonics 5. https://doi.org/10.1063/1.5140215.
Henson, S.A., Cael, B.B., Allen, S.R., Dutkiewicz, S., 2021. Future phytoplankton diversity in a changing climate. Nat. Commun. 12. https://doi.org/10.1038/s41467-021-25699-w.
Hieronymi, M., Bi, S., Müller, D., Schütt, E.M., Behr, D., Brockmann, C., Lebreton, C., Steinmetz, F., Stelzer, K., Vanhellemont, Q., 2023. Ocean color atmospheric correction methods in view of usability for different optical water types. Front. Mar. Sci. 10. https://doi.org/10.3389/fmars.2023.1129876.
Hillebrand, H., Dürselen, C., Kirschtel, D., Pollingher, U., Zohary, T., 1999. Biovolume calculation for pelagic and benthic microalgae. J. Phycol. 35, 403–424. https://doi.org/10.1046/j.1529-8817.1999.3520403.x.
Jaywant, S.A., Arif, K.M., 2024. Remote sensing techniques for water quality monitoring: a review. Sensors. https://doi.org/10.3390/s24248041.
Klöpper, S., John, U., Zingone, A., Mangoni, O., Kooistra, W.H.C.F., Cembella, A.D., 2013. Phylogeny and morphology of a Chattonella (Raphidophyceae) species from the Mediterranean Sea: what is C. Subsalsa? Eur. J. Phycol. 48, 79–92. https://doi.org/10.1080/09670262.2013.771412.
Kolodny, Y., Avrahami, Y., Zer, H., Frada, M.J., Paltiel, Y., Keren, N., 2022. Phycobilisome light-harvesting efficiency in natural populations of the marine cyanobacteria Synechococcus increases with depth. Commun. Biol. 5. https://doi.org/10.1038/s42003-022-03677-2.
Kudela, R., Berdalet, E., Enevoldsen, H., Pitcher, G., Raine, R., Urban, E., 2017. GEOHAB–the global ecology and oceanography of harmful algal blooms program: motivation, goals, and legacy. Oceanography 30, 12–21. https://doi.org/10.5670/oceanog.2017.106.
Kumar, B., Dikshit, O., Gupta, A., Singh, M.K., 2020. Feature extraction for hyperspectral image classification: a review. Int. J. Remote Sens. 41, 6248–6287. https://doi.org/10.1080/01431161.2020.1736732.
Kutser, T., 2004. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. Limnol. Oceanogr. 49, 2179–2189. https://doi.org/10.4319/lo.2004.49.6.2179.
Lavigne, H., Ruddick, K., Vanhellemont, Q., 2022. Monitoring of high biomass Phaeocystis globosa blooms in the southern North Sea by in situ and future spaceborne hyperspectral radiometry. Remote Sens. Environ. 282. https://doi.org/10.1016/j.rse.2022.113270.
Li, P., Ma, Q., Xu, S., Liu, W., Ma, Z., Ni, G., 2021. Opposite growth responses of alexandrium minutum and alexandrium catenella to photoperiods and temperatures. Plants 10. https://doi.org/10.3390/plants10061056.
Liu, Q., Trinder, J., Turner, I., 2016. A Comparison of sub-Pixel Mapping Methods for Coastal Areas. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences III–7, pp. 67–74. https://doi.org/10.5194/isprsannals-iii-7-67-2016.
Lodhi, V., Chakravarty, D., Mitra, P., 2019. Hyperspectral imaging system: development aspects and recent trends. Sens. Imag. 20. https://doi.org/10.1007/s11220-019-0257-8.
Luimstra, V.M., Verspagen, J.M.H., Xu, T., Schuurmans, J.M., Huisman, J., 2020. Changes in water color shift competition between phytoplankton species with contrasting light-harvesting strategies. Ecology 101. https://doi.org/10.1002/ecy.2951.
Lumini, A., Nanni, L., 2019. Deep learning and transfer learning features for plankton classification. Ecol. Inform. 51, 33–43. https://doi.org/10.1016/j.ecoinf.2019.02.007.
Maberly, S.Christopher, Gontero, Brigitte, 2022. Blue Planet, Red and Green Photosynthesis Productivity and Carbon Cycling in Aquatic Ecosystems. John Wiley & Sons. Incorporated.
Malhotra, A., Örmeci, B., 2023. Detection and identification of a mixed cyanobacteria and microalgae culture using derivative spectrophotometry. J. Photochem. Photobiol. B 238, 112616. https://doi.org/10.1016/j.jphotobiol.2022.112616.
Mao, Z., Stuart, V., Pan, D., Chen, J., Gong, F., Huang, H., Zhu, Q., 2010. Effects of phytoplankton species composition on absorption spectra and modeled hyperspectral reflectance. Ecol. Inform. 5, 359–366. https://doi.org/10.1016/j.ecoinf.2010.04.004.
Marampouti, C., Buma, A.G.J., Karin De Boer, M., 2021. Environmental Toxicity Assessment: State of the Art and Future Directions in a World of Arising Threats Mediterranean Alien Harmful Algal Blooms: Origins and Impacts. https://doi.org/10.1007/s11356-020-10383-1/Published.
McKibben, S.M., Schollaert Uz, S., Palacios, S.L., 2024. Testing a hyperspectral, bio-optical approach to identification of phytoplankton community composition in the Chesapeake Bay estuary. Earth and Space Sci. 11. https://doi.org/10.1029/2023EA003244.
Mehrubeoglu, M., Teng, M.Y., Zimba, P.V., 2014. Resolving mixed algal species in hyperspectral images. Sensors (Switzerland) 14. https://doi.org/10.3390/s140100001.
Méléder, V., Laviale, M., Jesus, B., Mouget, J.L., Lavaud, J., Kazemipour, F., Launeau, P., Barillé, L., 2013. In vivo estimation of pigment composition and optical absorption cross-section by spectroradiometry in four aquatic photosynthetic micro-organisms. J. Photochem. Photobiol. B 129, 115–124. https://doi.org/10.1016/j.jphotobiol.2013.10.005.
Millie, D.F., Schofield, O.M., Kirkpatrick, G.J., Johnsen, G., Tester, P.A., Vinyard, B.T., 1997. Detection of harmful algal blooms using photopigments and absorption signatures: a case study of the Florida red tide dinoflagellate, Gymnodinium breve. Limnol. Oceanogr. 42, 1240–1251. https://doi.org/10.4319/lo.1997.42.5_part_2.1240.
Millie, D.F., Schofield, O.M.E., Kirkpatrick, G.J., Johnsen, G., Evens, T.J., 2002. Using absorbance and fluorescence spectra to discriminate microalgae. Eur. J. Phycol. 37, 313–322. https://doi.org/10.1017/S0967026202003700.
Moloto, T.M., Thomalla, S.J., Smith, M.E., Martin, B., Louw, D.C., Koppelmann, R., 2023. Remote sensing of phytoplankton community composition in the northern Benguela upwelling system. Front. Mar. Sci. 10. https://doi.org/10.3389/fmars.2023.1118226.
Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.H., Devred, E., Bouman, H., 2008. Remote sensing of phytoplankton functional types. Remote Sens. Environ. 112, 3366–3375. https://doi.org/10.1016/j.rse.2008.01.021.
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 Sens. 13. https://doi.org/10.3390/rs13122381.
Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., Giardino, C., 2021. Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. Remote Sens. Environ. 253. https://doi.org/10.1016/j.rse.2020.112200.
Penna, A., Casabianca, S., Guerra, A.F., Vernesi, C., Scardi, M., 2017. Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. Mar. Genomics 36, 49–55. https://doi.org/10.1016/J.MARGEN.2017.06.001.
Pistocchi, R., Guerrini, F., Pezzolesi, L., Riccardi, M., Vanucci, S., Ciminiello, P., Dell'Aversano, C., Forino, M., Fattorusso, E., Tartaglione, L., Milandri, A., Pompei, M., Cangini, M., Pigozzi, S., Riccardi, E., 2012. Toxin levels and profiles in microalgae from the North-Western Adriatic Sea - 15 years of studies on cultured species. Mar. Drugs. https://doi.org/10.3390/md10010140.
Pokrzywinski, K., Morgan, C., Bourne, S., Reif, M., Matheson, K., Hammond, S., 2021. A Novel Laboratory Method for the Detection and Identification of Cyanobacteria using Hyperspectral Imaging: Hyperspectral Imaging for Cyanobacteria Detection. https://doi.org/10.21079/11681/40966.
Pyo, J.C., Duan, H., Baek, S., Kim, M.S., Jeon, T., Kwon, Y.S., Lee, H., Cho, K.H., 2019. A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery. Remote Sens. Environ. 233. https://doi.org/10.1016/j.rse.2019.111350.
Rocchini, D., Santos, M.J., Ustin, S.L., Féret, J.B., Asner, G.P., Beierkuhnlein, C., Dalponte, M., Feilhauer, H., Foody, G.M., Geller, G.N., Gillespie, T.W., He, K.S., Kleijn, D., Leitão, P.J., Malavasi, M., Moudrý, V., Müllerová, J., Nagendra, H., Normand, S., Ricotta, C., Schaepman, M.E., Schmidtlein, S., Skidmore, A.K., Šímová, P., Torresani, M., Townsend, P.A., Turner, W., Vihervaara, P., Wegmann, M., Lenoir, J., 2022. The spectral species concept in living color. J. Geophys. Res. Biogeosci. https://doi.org/10.1029/2022JG007026.
Roy, S., Llewellyn, C.A., Egeland, E.S., Johnsen, G. (Eds.), 2011. Phytoplankton Pigments. Cambridge University Press. https://doi.org/10.1017/CBO9780511732263.
Ruddick, K.G., De Vis, P., Goyens, C., Kuusk, J., Lavigne, H., Vanhellemont, Q., 2023. Second Derivative Water Reflectance Spectra for Phytoplankton Species Detection-Origin, Impact and Removal of Spectral Wiggles.
Salmi, P., Eskelinen, M.A., Leppänen, M.T., Pölönen, I., 2021. Rapid quantification of microalgae growth with hyperspectral camera and vegetation indices. Plants 10, 1–12. https://doi.org/10.3390/plants10020341.
Salmi, P., Calderini, M., Pääkkönen, S., Taipale, S., Pölönen, I., 2022. Assessment of microalgae species, biomass, and distribution from spectral images using a convolution neural network. J. Appl. Phycol. 34, 1565–1575. https://doi.org/10.1007/s10811-022-02735-w.
Santi, I., Kasapidis, P., Karakassis, I., Pitta, P., 2021. A comparison of DNA metabarcoding and microscopy methodologies for the study of aquatic microbial eukaryotes. Diversity (Basel). 13. https://doi.org/10.3390/d13050180.

Sathyendranath, S. (Ed.), 2014. Phytoplankton Functional Types from Space, no, 15. ed. IOCCG, Dartmouth, Canada.
Savitzky, Abraham, Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047.
Shen, A., Ishizaka, J., Yang, M., Ouyang, L., Yin, Y., Ma, Z., 2019. Changes in community structure and photosynthetic activities of total phytoplankton species during the growth, maintenance, and dissipation phases of a Prorocentrum donghaiense bloom. Harmful Algae 82, 35–43. https://doi.org/10.1016/j.hal.2018.12.007.
Sjöqvist, C., 2022. Evolution of phytoplankton as estimated from genetic diversity. J. Mar. Sci. Eng. https://doi.org/10.3390/jmse10040456.
Soja-Woźniak, M., Darecki, M., Wojtasiewicz, B., Bradtke, K., 2018. Laboratory measurements of remote sensing reflectance of selected phytoplankton species from the Baltic Sea. Oceanologia 60, 86–96. https://doi.org/10.1016/j.oceano.2017.08.001.
Solovchenko, A., 2023. Seeing good and bad: optical sensing of microalgal culture condition. Algal Res. https://doi.org/10.1016/j.algal.2023.103071.
Stelmakh, L., Kovrigina, N., Gorbunova, T., 2023. Phytoplankton seasonal dynamics under conditions of climate change and anthropogenic pollution in the Western coastal waters of the Black Sea (Sevastopol region). J. Mar. Sci. Eng. 11. https://doi.org/10.3390/jmse11030569.
Utermöhl, H., 1931. Neue Wege in der quantitativen Erfassung des Plankton.(Mit besonderer Berücksichtigung des Ultraplanktons.). SIL Proceed. 5, 567–596. https://doi.org/10.1080/03680770.1931.11898492, 1922-2010.
Vaddi, R., Phaneendra Kumar, B.L.N., Manoharan, P., Agilandeeswari, L., Sangeetha, V., 2024. Strategies for dimensionality reduction in hyperspectral remote sensing: a comprehensive overview. Egypt. J. Remote Sens. Space Sci. https://doi.org/10.1016/j.ejrs.2024.01.005.
Vaillancourt, R.D., 2004. Light backscattering properties of marine phytoplankton: relationships to cell size, chemical composition and taxonomy. J. Plankton Res. 26, 191–212. https://doi.org/10.1093/plankt/fbh012.
Valbi, E., Ricci, F., Capellacci, S., Casabianca, S., Scardi, M., Penna, A., 2019. A model predicting the PSP toxic dinoflagellate Alexandrium minutum occurrence in the coastal waters of the NW Adriatic Sea. Sci. Rep. 9. https://doi.org/10.1038/s41598-019-40664-w.
Wang, Z., Liu, L., Tang, Y., Li, A., Liu, C., Xie, C., Xiao, L., Lu, S., 2022. Phytoplankton community and HAB species in the South China Sea detected by morphological and metabarcoding approaches. Harmful Algae 118, 102297. https://doi.org/10.1016/j.hal.2022.102297.
Whitmire, A.L., Pegau, W.S., Karp-Boss, L., Boss, E., Cowles, T.J., 2010. Spectral backscattering properties of marine phytoplankton cultures. Opt. Express 18, 15073. https://doi.org/10.1364/OE.18.015073.
Wu, G., Ning, X., Hou, L., He, F., Zhang, H., Shankar, A., 2023. Three-dimensional Softmax mechanism guided bidirectional GRU networks for hyperspectral remote sensing image classification. Signal Process. 212. https://doi.org/10.1016/j.sigpro.2023.109151.
Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., Qiu, Z., 2015. Hyperspectral differentiation of phytoplankton taxonomic groups: a comparison between using remote sensing reflectance and absorption spectra. Remote Sens. 7, 14781–14805. https://doi.org/10.3390/rs71114781.
Xi, H., Hieronymi, M., Krasemann, H., Röttgers, R., 2017. Phytoplankton group identification using simulated and in situ hyperspectral remote sensing reflectance. Front. Mar. Sci. 4. https://doi.org/10.3389/fmars.2017.00272.
Xu, Z., Jiang, Y., Ji, J., Forsberg, E., Li, Y., He, S., 2020. Classification, identification, and growth stage estimation of microalgae based on transmission hyperspectral microscopic imaging and machine learning. Opt. Express 28, 30686. https://doi.org/10.1364/oe.406036.
Yang, M., Wang, W., Gao, Q., Zhao, C., Li, C., Yang, X., Li, J., Li, X., Cui, J., Zhang, L., Ji, Y., Geng, S., 2023. Automatic identification of harmful algae based on multiple convolutional neural networks and transfer learning. Environ. Sci. Pollut. Res. 30, 15311–15324. https://doi.org/10.1007/s11356-022-23280-6.
Yao, Y., Hu, C., Cannizzaro, J.P., Barnes, B.B., English, D.C., Xie, Y., Hubbard, K., Wang, M., 2023. Detection of Karenia brevis red tides on the West Florida shelf using VIIRS observations: accounting for spatial coherence with artificial intelligence. Remote Sens. Environ. 298. https://doi.org/10.1016/j.rse.2023.113833.
Zarco-Tejada, P., 2000. Chlorophyll fluorescence effects on vegetation apparent reflectance II. Laboratory and airborne canopy-level measurements with hyperspectral data. Remote Sens. Environ. 74, 596–608. https://doi.org/10.1016/S0034-4257(00)00149-8.
Zhang, S., Xiaoyun, L., Yuanyuan, F., Changlin, D., Yang, Y., Wang, J., Houjie, W., Jun, S., Shan, Z., Jing, W., 2016. Ecological provinces of spring phytoplankton in the Yellow Sea: species composition. Acta Oceanol. Sin. 35, 114–125. https://doi.org/10.1007/s13131.
Zhang, Y., Shen, F., Zhao, H., Sun, X., Zhu, Q., Li, M., 2024. Optical distinguishability of phytoplankton species and its implications for hyperspectral remote sensing discrimination potential. J. Sea Res. 202. https://doi.org/10.1016/j.seares.2024.102540.
Zhou, L., Wu, S., Gu, W., Wang, L., Wang, J., Gao, S., Wang, G., 2021. Photosynthesis acclimation under severely fluctuating light conditions allows faster growth of diatoms compared with dinoflagellates. BMC Plant Biol. 21. https://doi.org/10.1186/s12870-021-02902-0.
Zhou, X., Marani, M., Albertson, J.D., Silvestri, S., 2017. Hyperspectral and multispectral retrieval of suspended sediment in shallow coastal waters using semi-analytical and empirical methods. Remote Sens. 9. https://doi.org/10.3390/rs9040393.
Zhu, Q., Shen, F., Shang, P., Pan, Y., Li, M., 2019. Hyperspectral remote sensing of phytoplankton species composition based on transfer learning. Remote Sens. 11. https://doi.org/10.3390/rs11172001.
Zingone, A., Escalera, L., Aligizaki, K., Fernández-Tejedor, M., Ismael, A., Montresor, M., Mozetič, P., Taş, S., Totti, C., 2021. Toxic marine microalgae and noxious blooms in the Mediterranean Sea: a contribution to the global HAB status report. Harmful Algae 102. https://doi.org/10.1016/j.hal.2020.101843.