Silvia Elisabeth Zieger, MSc

Multiparametric Sensing –

from Development to Application in Marine Science

Dissertation

Zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

eingereicht an der

Technischen Universität Graz

Betreuer:

Univ.-Prof. Dipl.-Chem. Dr.rer.nat. Ingo Klimant

Institut für Analytische Chemie und Lebensmittelchemie

Graz, May 2019

"I may not have gone where I intended to go, but I think I have ended up where I needed to be." D. Adams

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

_____

Date

_____

Signature

Abstract

Monitoring and control of coastal environments has gained significant importance since they are subject to increased exploitation by recreational facilities and aquacultures. Marine environments are vulnerable ecosystems and their dynamic and fragile equilibrium is compromised by diverse anthropogenic and natural pollutants. However, since we know little about the related biological degradation processes and chemical dynamics at the time, appropriate sensor observation networks targeting the determination of relevant biochemical parameters are crucial for elucidation and enhancing our knowledge.

Facing these challenges, we developed a concept for a multiparametric optical sensor system enabling the real-time identification of phytoplankton at unicellular level and the simultaneous determination of relevant metabolic parameters, such as pH, oxygen, carbon dioxide or temperature. In the first step, a compact multi-channel algae sensor was realized within the EU-project SCHeMA. The algae sensor features internal calibration and standardization strategies for inter- and intra-comparability. Since the algae sensor consists of readily available and low-cost optical as well as electronic components, the sensor is an attractive monitoring tool especially for small agencies and research institutions. To enable a real-time identification on site, the sensor is combined with a multivariate pattern recognition algorithm that enables the identification of the phytoplankton with respect to their characteristic pigment pattern. The system performance of the algae sensor was validated via a leave-one-out cross validation using 53 different unialgal cultures resulting in a separation efficiency and an overall accuracy of 97.3%. Its ease of use as well as its versatility were demonstrated during multiple international field trips in Italy and France.

Additionally, concepts for multiplexed optical sensing were developed based on the modified dual lifetime referencing (DLR) technique combining two indicators sensitive to different analytes of interest with similar spectral properties but largely different luminescence decay times. However, due to the assumptions made during data evaluation, hybrid sensors are prone to error which has to be assessed by inclusion of an error estimation and the investigation of its propagation across the evaluation process. Thus, the integration of these optical hybrid sensors into the algae sensor would yield a versatile and compact multiparametric sensor that poses a valuable contribution to existing monitoring tools in marine science. Its possibility to determine multiple analytes at the exact same position enables a high spatial resolution and its conservation of resources makes it an attractive option, in particular, for small agencies and universities.

Kurzfassung

Seit Küstengebiete vermehrt von Erholungseinrichtungen und Aquakulturen genutzt werden, hat ihre Überwachung und Kontrolle deutlich an Bedeutung gewonnen. Darüber hinaus sind die Meere gefährdete Ökosysteme und ihr dynamisches und fragiles Gleichgewicht ist durch verschiedene anthropogene und natürliche Schadstoffen beeinträchtigt. Da wir allerdings derzeit wenig über ihre entsprechenden biologischen Abbauprozese und chemischen Dynamiken wissen, sind adäquate Überwachungsnetzwerke, die auf die Bestimmung relevanter biochemischer Parameter abzielen, essentiell um neue Erkenntnisse zu gewinnen.

Angesichts dieser Herausforderungen entwickelten wir ein Konzept für ein mehrparametrisches optisches Sensorsystem, welches eine Echtzeit-Identifizierung des Phytoplanktons auf unizellularem Level ermöglicht und gleichzeitig die Bestimmung relevanter metabolischer Parameter wie pH, Sauerstoff, Kohlenstoffdioxid oder Temperatur, erlaubt. Im Rahmen des EU-Projektes SCHeMA wurde hierfür im ersten Schritt ein kompakter mehrkanaliger Algensensor entwickelt. Den Sensor zeichnen insbesondere die internen Kalibrations- und Standardisierungs-strategien aus, welche die Vergleichbarkeit zwischen den einzelnen Kanälen sowie zwischen verschiedenen Geräten gewährleistet. Da der Algensensor aus gängigen und preiswerten optischen und elektronischen Komponenten aufgebaut ist, stellt er besonders für kleinere Behörden und Forschungseinrichtungen ein attraktives Kontrollinstrument dar. Um eine Echtzeit-Identifizierung vor Ort zu ermöglichen, ist der Sensor mit einem multivariaten Mustererkennungsalgorithmus ausgestattet, der die Identifizierung des Phytoplanktons basierend auf seiner inhärenten Pigmentstruktur ermöglicht. Die Leistungsfähigkeit des Algensensors wurde in einem Kreuzvalidierungsverfahren mit 53 verschiedenen artenreinen Kulturen überprüft, wobei eine Trenneffizienz und eine Gesamtgenauigkeit von 97,3% erzielt wurden. Der Bedienkomfort und ebenso seine Einsatzflexibilität wurden in mehreren internationalen Fieldtrips in Italien und Frankreich bewiesen.

Ergänzend dazu wurden, basierend auf der Idee der modifizierten dual lifetime referencing (DLR) Methode, Konzepte für optische Multiplex-Sensoren entwickelt. Hierbei wurden zwei Indikatoren kombiniert, die sensitiv auf verschiedene Analyten sind und ähnliche spektrale Eigenschaften, aber deutlich voneinander abweichende Lumineszenz-Lebenszeiten aufweisen. Aufgrund der hohen Komplexität, sind Hybridsensoren jedoch fehleranfällig. Durch die Berücksichtigung einer Fehlerabschätzung und der Fehlerfortpflanzung über den Evaluationsprozess hinweg, konnte die Fehleranfälligkeit reduziert werden. Demzufolge würde die Eingliederung der optischen Hybridsensoren in den Algensensor zu einem vielseitig einsetzbaren und kompakten multiparametrischen Sensor führen, der als eine wertvolle Unterstützung bereits existierender Küstenüberwachungsinstrumente angesehen werden kann. Seine Fähigkeit mehrere Analyten an exakt der gleichen Position zu bestimmen, ermöglicht dabei eine hohe räumliche Auflösung und

sein ressourcenschonendes Design machen den Multiplex-Sensor zu einer interessanten Lösung besonders für kleine Büros und Universitäten.

Dedication

Now that the time of leaving Graz is approaching, I want to thank the people with whom I was fortunate enough to go part of my way. Let it be people who I have met only for a short moment or let it be people who have accompanied me over long distances, I am grateful to all of you.

In particular, I would like to thank you, Ingo, not only for being my supervisor and giving me the opportunity to continue working on my satellite project. Many thanks also for your confidence in me and my work that I would cope with the given challenges. Thank you for encouraging me in times of doubt and teaching me caution whenever I was restless.

Gü, you are one of the persons, I met only briefly, but you inspired and encouraged me a lot and I am pleased that I had the possibility to learn from you. Many thanks for your faith and your assistance at any time, even if you had a great deal to do. Maybe someday, we will work together – you never know. It would be a pleasure!

The hardest thing by far is to find words your all of you at the ACFC. All of you contributed to the fact that the time at the institute will be unforgettable and I am pleased that I could spend some time with you – during and after work. I can hardly imagine how it will be without you. A special thanks goes to you, Sergey! Thanks for your support and your time, when I finally got started with the optical sensors. A big 'thank you' is for my office mates! Fips, thanks for the feeling of belonging. It was a pleasure to spend time with you, discussing, chatting and playing Kamisado (→ KaKaDo). It's amazing how you handle every situation and you can be proud of your scientific and personal achievements. Thank you Max for your attentiveness and your casualness, when things didn't work out as expected. One of those things, I can learn from you apart from the Bavarian dialect. Irene, you are just amazing – it is awesome, how you manage all those projects and how you still can joke and take an interest in others and their problems. Another thanks go to BerniJ and Christoph for your assistance and support whenever I struggled or when I crashed Python. Many thanks for all the effort you put into the correction of this thesis. I am not sure where I could find better early-morning coffee buddies! Many thanks also to you Heidi, for always wearing a smile no matter how stressful it is and for spontaneous wine evenings! A special thank goes to my only student whom I had the chance to supervise to. Thanks Anna K for your staying power during monotonous tasks and for not losing your sense of humor in hairy situations ('If you do not mind – I guess the autoclave is burning.'). Also a big thank-you to my co-worker, Lukas, for your ingenuity and the numerous discussions. Without you, the ALPACA would never have been possible. Iris, although we have not known each other for a long time, I appreciate all the adventures we share so far and I am looking forward to the next experiences we will gain together! Many thanks to you Peter, for the joint walks but not for the voice mails that cannot substitute a conversation face to face. A big thank you to you, Tanja, and your capability to take delight in the minutiae of everyday life. David, I honor

Content

Part I

Introduction

1. Scope and Outline of the Thesis

Marine environments are vulnerable ecosystems facing diverse anthropogenic constituents, natural substances and organisms that might have an adverse effect on their fragile equilibrium, their marine life and on human health. To preserve marine ecosystems, it is thus crucial to understand biological processes in marine environments and further to control relevant types of hazards and significant quality parameters at an appropriate temporal and spatial scale. The early-warning, monitoring and control of these constituents and organisms enables finally the evaluation and management of their risks and consequences for marine ecosystems.

Although there exist already a huge number of different monitoring tools enabling the control and management of coastal areas, these systems are often expensive, bulky and inconvenient in their utilization. In recent years, the trend also goes towards sensor observation networks enabling the simultaneous determination of multiple parameters at the exact same site. Thus, the scope of this doctoral thesis was to address this issue and to develop a novel multiparametric sensor system targeting the simultaneous control of different marine quality parameters in one miniaturized and low-cost monitoring tool.

This thesis is written as comprehensive work of all projects that were dealt with during the past three years of the PhD and consists of four different parts. The first part is dedicated to the theoretical background of the algae monitoring and the multiparametric sensing, while the second and third part are dedicated to the outcome of those projects. The last part however, provides a comprehensive conclusion of the dissertation and combines the individual projects to a coherent topic. All manuscripts presented are already accepted and published in peer reviewed journals. Other projects are supposed to be published in the near future.

Chapter 2 introduces the necessary theoretical background on algae, provides an overview of available monitoring tools and discusses their advantages and challenges. It emphasizes the requirements of a convenient multiparametric observation sensor and explains the multivariate discriminant analysis implemented for correct evaluation of measured data. Moreover, in this chapter, smart calibration and standardization strategies are explained that are crucial for long-term deployment of the sensor.

Chapter 3 presents the necessary background on multiparametric sensing using optical hybrid sensors and the significance of error estimation as well as the propagation of errors during evaluation. Furthermore, this chapter provides a general description of the mathematical principles required for multiparametric sensing and its implementation in four different optical hybrid sensors.

Chapter 4 describes the development and the fabrication of the algae sensor optimized for a long-term deployment in coastal areas and the real-time monitoring of algae dynamics. In particular,

this chapter focuses on the technical specifications of the sensor, its operation and the implemented calibration strategies to ensure the inter- and intra-comparability of the sensor. Its performance was validated during a long-term field campaign at the port of Genoa (Italy).

Chapter 5 describes the operating principle of the linear discriminant algorithm as preferred evaluation strategy of the algae sensor. The performance of the algorithm was tested as well as the influence of different cultivation factors on its separation effectivity. In addition, a comprehensive analysis of the pigment composition and their spectral properties are presented.

Chapter 6 contains the operating instructions for the graphical user interface that was programmed in Python for real-time analysis and evaluation of the measurements conducted by the algae sensor. Besides the installation, the software interface and all features are explained in detail.

Chapter 7 includes additional experiments that were conducted by the algae sensor, in particular, this chapter presents a comprehensive overview of three international field trips. The algae sensor was deployed in Genoa (Italy), Arcachon (France) and Gironde (France) for long-term stability tests and for performance tests under environmental conditions as well as to check the compatibility of the sensor within a larger sensor network.

Chapter 8 contains the operating instructions for the dualSense software that was programmed in Python targeting the evaluation of multiplexed sensors. The software interface is explained in detail as well as the application of the software on different optical hybrid sensors with special focus on the required calibrations of the individual hybrid sensors.

Chapter 9 provides preliminary experiments of the single-layered dual optical sensor sensitive for oxygen and temperature applied in the gaseous phase. Besides the measurement setup, the calibration of the sensor is described enabling the extrapolation of the discontinuous measurement data to a reference matrix for simultaneous determination of oxygen and temperature. The initial goal to implement an appropriate calibration strategy and to characterize the single-layered sensor with respect to both analytes was successfully realized.

2. Marine Algae and Coastal Monitoring

This chapter introduces the necessary theoretical background of micro algae and cyanobacteria, provides an overview of available monitoring tools and discusses the requirements for real-time identification and early-warning systems. It emphasizes the chosen approach of a multivariate pattern recognition algorithm which is combined with a classification algorithm based on the Mahalanobis distance.

2.1. Algae Monitoring

2.1.1. Algae

2.1.1.1. Definition of Algae

The group of algae encompasses a heterogeneous and diverse collection of different autotrophic organisms widely spread all over the world occurring both in terrestrial environments and in water bodies. Species of algae are extremely diverse in their morphology and range in size from few micrometers to several meters long. The most common definition of algae is given by Lee in 1980:

"The algae are thallophytes (plants lacking roots, stems, and leaves) that have chlorophyll-a as their primary photosynthetic pigment and lack a sterile covering of cells around the reproductive cells."[1]

With this wide definition also cyanobacteria, formerly known as blue-green algae, are included, although they might be closer related in evolution to prokaryotic bacteria than to the eukaryotic algae. Both, algae and cyanobacteria, are one of the basic food sources in the food web, as they are major producers of organic matters and oxygen necessary for the consumer's metabolism. Their position in the food web combined with their ability to respond rapidly and predictably to a range of pollutants make them an interesting and sensitive indicator for changing environmental conditions.[2] In contrast to their capability to provide early-warning signals for changing conditions, some algal species itself can have an adverse impact on their surroundings. In addition, the occurrence of algal events as well as their intensity have increased and a significant increase of economical exploitations in coastal environments has been observed at the same time.[3,4] Thus, it is worth to study the occurrence of algae especially in coastal environments and their bloom dynamics including their bloom composition and related biological processes at a temporal and spatial scale.[5]

As an initial step, this doctoral thesis focuses on the early-stage identification of cyanobacteria and dinoflagellates in algal blooms of mixed compositions as both organisms are well-known to pose a severe threat to marine livings and human health.

2.1.1.2. Taxonomic classification

Algae can be classified according to their morphology into hierarchical structures, similar as it is done for other plants and animals. This concept was first invented by Carl Linnaeus in 1753, however, concepts change during decades.[6] Nowadays, taxonomic classification is done based on common morphological and evolutionary characteristics. Algae with shared characteristics belong to the same *taxa*, which is further classified into more specialized *ranks*. A brief overview of the general taxonomic scheme is provided in Figure 2-1.
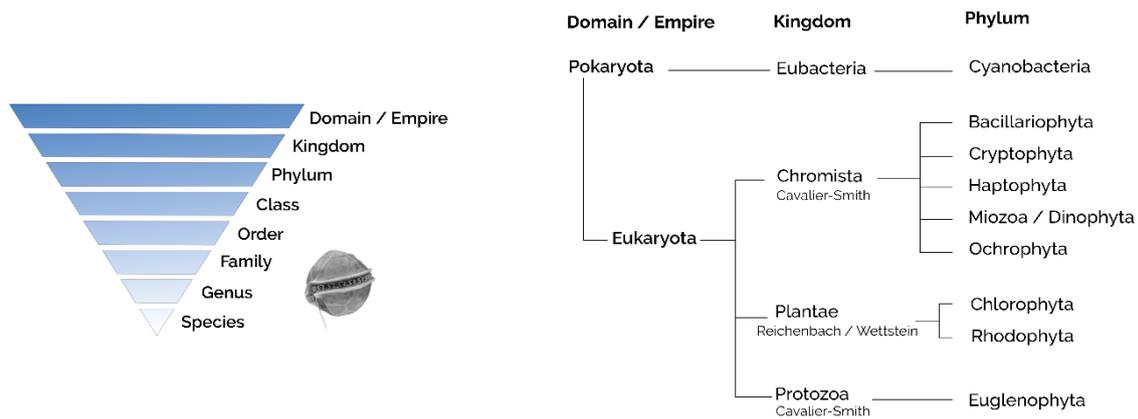
Figure 2-1: General taxonomic classification based on the systems introduced by Cavalier-Smith[7], Reichenbach[8] and Wettstein[9] (**left**) and overview of the taxonomic classification of algal groups used in this doctoral thesis (**right**).

Within this thesis, which is embedded in the working framework of the EU funded SCHeMA-project (FP7-OCEAN 2013.2-Grant Agreement 614002), only marine and unicellular microalgae occurring in coastal environments were analyzed. These marine algae, ranging in their cell size between few pico- to micrometers, are called phytoplankton. In Figure 2-1 all nine algal classes studied during this project and their taxonomic classification are provided. For further details on their taxonomy, please refer to the supporting information of chapter 4 and chapter 5.

2.1.2. Significance of Algae Monitoring

In a strict sense, harmful algal blooms (HABs), commonly known as 'red tides' are natural phenomena of rapidly growing microalgae that might have an adverse effect to their surrounding environment, public health or to economic and ecological structures.[10] In general, this definition of HABs encompasses a wide range of different phenomena, caused by nontoxic and toxic microalgae that cover coastal environments all over the world, leading to unfavorable conditions for marine livings and humans. The reason for this wide definition is the heterogeneity and diversity of microalgae responsible for these phenomena and that the phenomena itself are heterogeneous and diverse.[11]

On the one side, nontoxic HABs can cause harm due to the biomass they achieve during growth, as they can clog gills or lungs of fishes and marine mammals. Moreover, their dense biomass changes the light penetration into the water and the oxygen content within the water body during respiration as well as during bloom depletion, leading to a low oxygen content or even to hypoxic dead zones.[12–14] On the other side, toxic algal blooms cause harm to their surrounding environment and to human health by endogenous biotoxins they produce and which can accumulate in the food web. After accumulation, these phycotoxins can kill fish, oysters and marine mammals, leading to tremendous economic losses for fish farms, aquacultures and coastal recreational facilities.[3,15] Besides, these toxins can affect human health after ingestion of contaminated fish and oysters or after drinking contaminated water, leading to various disorders, known as shellfish poisoning, or even to death. The most important representatives of these harmful algae are dinoflagellates and cyanobacteria, formerly known as blue-green algae and both of them are widely spread all over the world.

In recent decades, the frequency, intensity and evident geographical distribution of bloom events have increased all over the world. Whereas previously only few countries have been facing problems with algae, currently most coastal environments in the world are affected by harmful algal blooms according to Anderson.[16] The reasons for this trend are discussed differently ranging from natural phenomena such as dynamic distribution processes and circulation of algae to human related phenomena such as unusual climatological conditions due to climate change or pollution-related nutrient enrichment. Furthermore, the increased utilization of coastal waters for drinking water treatment and aquacultures increases the perception of such bloom events and their risks for marine livings and humans health.[3,11,16,17] Due to their evident increase and their potential risk even at low concentrations, the awareness of marine scientists for HABs has grown together with their demand to enhance our understanding of bloom dynamics and related biological processes.

2.1.3. State-of-the-Art Monitoring Systems

The early-stage identification and prediction of harmful algal blooms rely on observations to monitor and control algal dynamics and their distributions in relation to environmental conditions and models to interconnect these observations.[10] Moreover, since algal dynamics and bloom distributions are affected by, and, in turn, affect their surrounding environments, a comprehensive observation of multiple parameters is required. In addition, since algal blooms are patchy and episodic, technologies are required that track algal distributions on a temporal and spatial scale. Although both criteria are important, a synoptic observation of large areas as well as a precise control of small fixed points, both cannot be achieved at the same time with the same instrument. Nevertheless, to address both subjects, instruments developed for the monitoring of algae are always part of a larger surveillance program. These observing programs are cooperative networks combining different technologies, methods and disciplines for synergetic surveillance of coastal areas in order to cover a wide temporal and spatial range and to improve our understanding of

complex biological processes in a dynamic environment. Examples for such joint European or international observation programs are JERICO next (jerico-ri.eu)[18], JPI Ocean (jpi-oceans.eu), Copernicus (marine.copernicus.eu), the Ocean Networks Canada (ONC, oceannetworks.ca) and the Ocean Observatories Initiative (OOI, oceanobservatories.org). However, the review of all available monitoring systems is far beyond the scope of this thesis, wherefore this thesis provides a brief overview of different approaches using optical methods for the algal monitoring. For further detail on available monitoring technologies, reference is made to the reviews of Anderson, Zielinski and Moore.[10,15,19]

2.1.3.1. Requirements for Algae Monitoring

The requirements for algal monitoring instruments strongly depend on their final application, however, since algal events are episodic and patchy, continuous and in situ detection is preferred instead of time, energy and cost-intensive laboratory work, which may miss episodic algal events. Moreover, real-time identification of bloom events and their compositions are crucial for data assimilation and the forecast of algal events.[20] Consequently, the system must be sensitive and selective enabling the early-stage identification of algae at very low concentrations before they can cause any adverse effect on their surroundings. Since toxic algae usually compose the minor part in algal blooms of mixed composition and their cell size might vary widely from few picometers to some centimeters or even meters, their distinct detection and monitoring is even more challenging. Therefore, appropriate instruments are required for reliable and precise quantification and classification of algal blooms.[10,16,17]

Besides, algal monitoring instruments ought to be flexible for different applications and for easy integration onto different platforms, enabling a synoptic observation for a comprehensive monitoring and management of algal blooms across coastal environments. Thus, a compact and miniaturized design is preferred for the application on small (autonomous) devices (e.g. autonomous underwater vehicles (AUVs), gliders, Argo floats and surface moorings, etc.). In addition, the deployment of the sensor and its maintenance ought to be simplified without extended technical and laboratory effort or the demand for highly trained personnel. In order to ensure a user-friendly handling and a long-term stability of the instrument, the system might be embedded into a waterproofed enclosing and the whole system should either consume only little energy for its maintenance or provide an independent power supply. Besides, the algae monitoring tools ought to be stable and robust over time for long deployments without re-calibration or drifts due to biofouling or other interferences and artifacts. In particular, the prevention of optical instruments from biofouling and other artifacts, such as colored dissolved organic matter (CDOM), hydrocarbon materials or optical brighteners, plays a central role in the development of in situ monitoring devices.[15] For the final application in transnational and international observation programs, the monitoring tools, in particular, those which target the monitoring of algal blooms on an intermediate or small scale, must be widely spread across a large area and in large quantities.

Therefore, the system and its components ought to be low-cost and affordable for small research institutions and fisheries.

For real-time observation and early warning of rapidly changing conditions, the instrument should work continuously and subsequently transmit the results wirelessly to the marine station. Moreover, for additional and independent data backup, the sensor should be able to log data directly on the device. Besides the data transmission and storage, a high sampling rate is required to map patchy and rapidly changing conditions and alterations in marine life on a high temporal resolution.

2.1.3.2. Optical Monitoring Tools for Algal Monitoring

Various approaches for optical algal monitoring targeting the observation on a wide temporal and spatial scale are under development or commercially available on the market.[19,21]

To cope with the requirements for a real-time monitoring of algae, joint sensory observing networks are established all over the world in order to collate information about physical and biological processes affecting the phytoplankton ecology and their bloom dynamics, as it was mentioned before.[20,21] An overview of these biological and physical processes as well as the sampling domain of various monitoring platforms is provided in Figure 2-2.
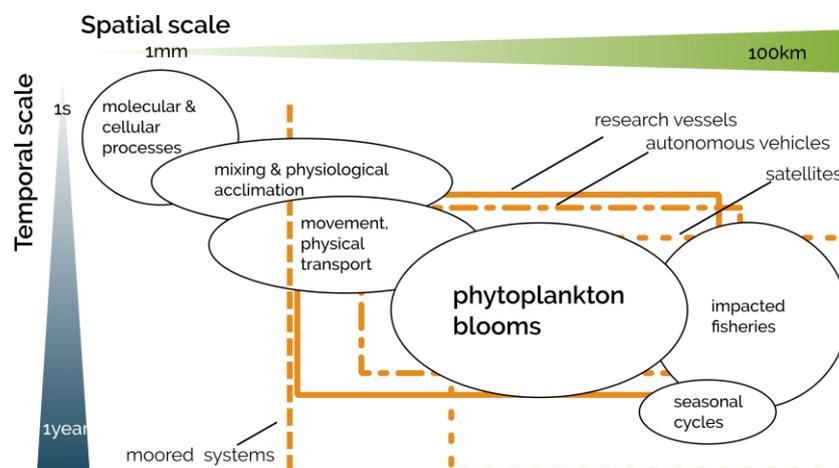
Figure 2-2: Comprehensive overview of biological and physical processes effecting the phytoplankton ecology and the sampling domains of various monitoring platforms (squares). Redrawn from Schofield.et al.[20]

For synoptic observation across large areas, remote sensing from satellites or aircrafts is the preferred system, that detects and identifies phytoplankton dynamics from the distance.[10,19] However, since light is attenuated and absorbed when penetrating the water surface, satellites and airborne systems are limited to the analysis of surface waters. Moreover, the radiative transfer within the atmosphere and, consequently, their dependency on varying weather conditions are further challenges of remote sensing systems.[19] In general, remote sensing requires the change in the absorbance, scattering or reflective properties of the surface water at certain wavelengths, and

thus high biomass concentrations are required to affect the color of the ocean. Since toxic algal species usually pose the minor part of algal assemblages, remote sensing systems are not able to discriminate toxic species from non-toxic species in mixed assemblages. In addition, remote sensing is also limited in its temporal resolution. In the case of satellite-born systems, the time for re-visiting the observing area depends on its orbital velocity and is usually several days. In the case of airborne systems, the temporal resolution is also low, as the sensing mission is often carried out as one-time operation.[10]

For real-time monitoring and early-stage identification and prediction of (harmful) algae, however, continuous measurement on an intermediate to small spatial scale with high temporal resolution is required. There is a high potential for highly accurate and precise instruments targeting the sampling domain, where synoptic observation systems and high resolution point detection instruments intersect.[15,19] Various approaches for direct or indirect observation of phytoplankton dynamics exist and they are either already commercially available on the market or under development. Indirect observation systems target the detection of changing physical, biological or chemical parameters, such as the oxygen content or the turbidity of surface water. Indirect measurements of such parameters are often simple and no complex calibration procedure is required. Moreover, their ease of use is another argument for their broad-scale use by end-users and untrained personnel. However, indirect observations of physical, biological or chemical parameters, such as turbidity, can only study trends of total contents instead of high-resolution and detailed analysis of phytoplankton dynamics. In particular, the control of turbidity by the Secchi disk or other automated turbidity instruments integrated into CTD sensors can only record the total content of all dissolved particles, organisms and compounds attenuating the clarity of the surface water.[10] Therefore, direct measurement and control of phytoplankton dynamics are preferred to study the phytoplankton ecology. Different approaches for the direct observation exist – ranging from the image-based identification of the cell shape and morphology to the pigment-based classification of algae according to certain light-harvesting pigments. State-of-the-art technologies for direct observation of the phytoplankton ecology include the HPLC analysis and the microscopic analysis. However both are restricted to laboratory applications and they are tedious and require highly trained personnel.[10,22] Thus, these approaches are limited in their temporal and spatial coverage and are inappropriate for real-time identification and early-warning of episodic or rapidly changing bloom events.[20]

In general, spectrofluorometric instruments are preferred for real-time and in situ monitoring compared to spectrophotometric instruments, although in situ applications based on the evaluation of absorption coefficients exist in the literature or are commercially available.[20,23−27] Both methods are simple in their handling, affordable, non-invasive and suitable for miniaturization. However, spectrofluorometric instruments are considerably more sensitive and selective than spectrophotometric methods.[10] In particular, spectrophotometric methods are prone to errors due to interferences coming from scattering particles or instabilities on the instrument side, such as an intensity drift of the light source. On the other side, even spectrofluorometric methods suffer from uncertainties due to photoacclimation, physiological state of the phytoplankton and other

ecological and environmental aspects. For further details on the chlorophyll fluorescence and its dependencies, please refer to section 2.2. As a result, internal standardization strategies and carefully calibration of the instruments are crucial for reliable classification and quantification of algae assemblages.[15,19,20]

Focusing on spectrofluorometric methods, (imaging) flow cytometry is definitely prevailing as it combines the spectrofluorometric analysis at certain wavelengths with morphological evaluations of individual cells. Although state-of-the-art technologies are usually constrained to the laboratory, in situ versions are commercially available, such as the CytoBuoy, the CytoBot or the Imaging FlowCytobot.[28–31] These instruments have proven to be valuable and efficient in discriminating and quantifying phytoplankton assemblages, even though they are often expensive and time- and energy-consuming, wherefore they are not suitable for continuous monitoring widely spread across coastal areas. The miniaturization of these technologies has been object of research and especially the working group of Francis Ligler realized the implementation of a microflow cytometer in autonomous underwater vehicles (AUVs).[32–34] Although, these approaches are suitable for in situ monitoring, they partly diminish their outstanding performance and filamentous cells can hardly be analyzed within the microfluidic setup. Thus, microflow cytometers are primarily applied in medicinal point-of-care examinations rather than in marine science.[33]

Simplified sensor systems, such as single channel fluorometers, targeting the in situ quantification of overall parameters, are commercially available and comprehensively evaluated by the ACT, a partnership of research institutions, managers and private sector companies.[35] Overall parameters are colored dissolved organic materials, other anthropogenic materials, chlorophyll or additional photopigments, such as phycoerythrin and phycocyanin.[19] While these single channel fluorometers are effective in quantifying the approximate phytoplankton concentration, they often suffer from bio-fouling and have to cope with uncertainties due to environmental conditions as mentioned before.[19] Moreover, they are constrained to the discrimination of algae and cyanobacteria.

More complex spectral fluorometers using different excitation and emission bands provide a greater potential for the discrimination of phytoplankton assemblages and interfering dyes. They reveal taxonomic information after evaluation of the relative composition of marker pigments and they can provide information about the photosynthetic activity, when the fluorescence emitted from both primary photosystems in compared.[19] Over the past decades, the component technology has been evolved and few products are now commercially available, e.g. the FluoroProbe and the AlgaeOnlineAnalyser from bbe Moldaenke (bbe-moldaenke.de) or the Cyclops-7 from Turner Designs (turnerdesigns.com). However, the optimization of these technologies and, in particular, their strategies to deal with problems, such as anti-fouling or photoadaption of living organisms, are subjects of research. In addition to the technical development, strategies for signal processing and data interpretation are required and crucial to deconvolute reliable information on the phytoplankton composition in mixed bloom compositions. Various approaches using multivariate analyses have been emerged over time, ranging from the analysis of the fourth-derivative of

absorbance spectra coupled with a similarity algorithm to statistical classification methods, such as the factor analysis or the linear/quadratic discriminant analysis.[36–38]

In conclusion, although a few instruments already exist, there is a high potential for upcoming technologies and deconvolution software targeting the real-time identification of phytoplankton groups in blooms of mixed compositions at an intermediate scale with a high temporal resolution. Challenges, such as interfering colored dissolved organic particles and other anthropogenic materials, and anti-fouling methods have to be addressed. Internal calibration and standardization strategies are required for enhanced inter- and intra-comparability within and between different technologies. In order to contribute significantly to existing monitoring systems, the developed sensors ought to be affordable for small companies and research institutions as well as suitable for flexible applications on board or integrated into submersible housings.

2.1.3.3. ALPACA

To address these issues, a miniaturized sensor module using multiple excitation and two emission bands was invented within the framework of the EU-funded SCHeMA-project (FP7-OCEAN 2013.2 - Grant agreement No. 614002). The module is called ALPACA – Advanced Luminescence-based Phytoplankton Analysis and Classification Appliance. The ALPACA is constructed in a flow through mode combining advantages of a flow cytometer and a spectrofluorometer. Since readily available and low-cost optical and electronic compounds are incorporated into the ALPACA, the sensor module is affordable for small companies and research institutions. For appropriate and reliable deconvolution of the recorded signal, a multivariate analysis for pattern recognition is combined with a classification algorithm based on the Mahalanobis distance. For detailed information on the technical and optical setup as well as the deconvolution algorithm, references are made to chapter 4 and chapter 5. It has to be emphasized that the ALPACA contributes significantly to existing observation systems, however, few limitations remain. The effects of photoadaption and varying pigmentation due to environmental conditions, stressors such as restricted nutrient availability, pesticides or other pollutants, or due to the stage of growth have to be investigated further.

2.2. Photosynthesis and Chlorophyll Fluorescence

2.2.1. Brief Overview of the Photosynthesis

Photosynthesis is a process performed by photoautotrophic organisms such as plants, algae and some bacteria to build up energy-rich organic compounds from inorganic compounds and water after capturing energy from the sunlight and conversion of the solar energy into chemical energy. All phototrophic algae and oxyphotobacteria use carbon dioxide as oxidant and produce oxygen during photosynthesis, wherefore this process is called oxygenic photosynthesis. However, if other reductants are used, this process is called anoxygenic photosynthesis.

Photosynthesis is an essential part of the food web as most organisms rely directly or indirectly on this process for nourishment: On the one hand, organisms rely on photosynthesis when they use this process to maintain their energy household and to feed themselves. On the other hand, other organisms consume photoautotrophic algae and bacteria and use the oxygen evolved for (cellular) respiration.
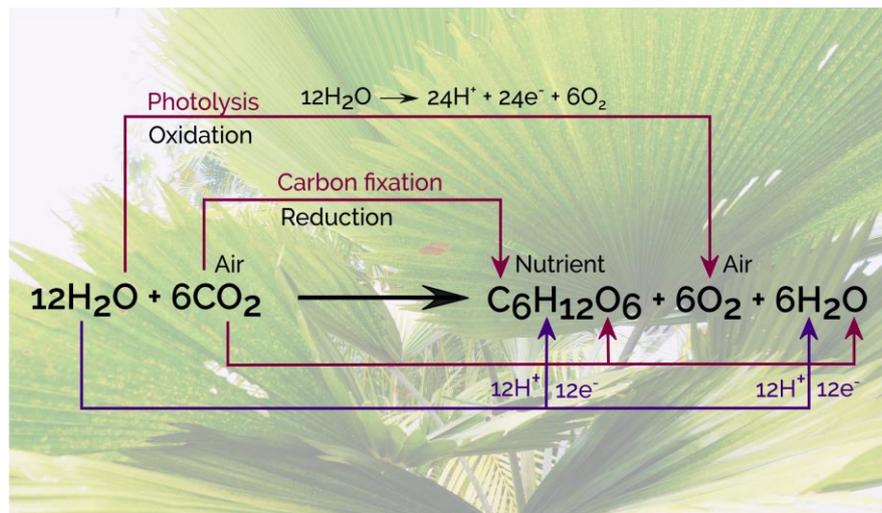
Figure 2-3: General scheme of photosynthesis performed by algae and oxyphotobacteria. During the light reaction, 24 protons are generated via photolysis after capturing of solar energy and conversion of the energy into chemical energy. Protons are used to drive the carbon dioxide reduction taking place during the Calvin cycle.

An overview of the overall reaction of the photosynthesis is provided in Figure 2-3. The whole process consists of two partial reactions conducted in different parts of the cells. The first part, located at and in the thylakoid membrane, is called light reaction as this part is light dependent and the solar energy is captured by several photopigments and subsequently converted into chemical energy. This energy is used for oxidative photolysis of water or other reducing agents in order to drive the synthesis of ATP and NADPH via a complex electron transport chain. ATP and NADPH are high-energy compounds that are required to store electrons and protons, which are generated during photolysis. In the following dark reaction, or Calvin cycle, ATP and NADPH are subsequently

used to synthesize energy-rich organic compounds after fixation of atmospheric carbon dioxide and evolution of oxygen as bi-product. In algae and higher plants, the dark reaction is performed in the stroma of the chloroplasts. However, since cyanobacteria lack any membrane-bound organelles, their thylakoid membrane is spread out within the cell interior. [39,40]

The whole process of the photosynthesis is sensitive and adjustable to changing environmental conditions, including actual and elapsed irradiation or the nutritional state of the cells, in order to maintain the cellular integrity. Essential parts to control and protect cells from photodamage are either quenching reactions implemented in the Calvin cycle or the photosystems implemented into the light reaction. These photosystems convert solar energy into chemical energy and therefore maintain the electron transport chain during light reaction. They further facilitate various protective and control mechanism to cope with unfavorable environmental conditions.

2.2.2. Light Reaction and Photosystems

Since the light reaction contains all information required for the classification algorithm and further description of the subsequent dark reactions are beyond the scope of the thesis, only the light reaction is presented here with special focus on the photosystems used for light conversion. However, for detailed information about the biochemical background of the dark reaction and the photosynthesis in general, reference is made to Govindjee et al.[41]
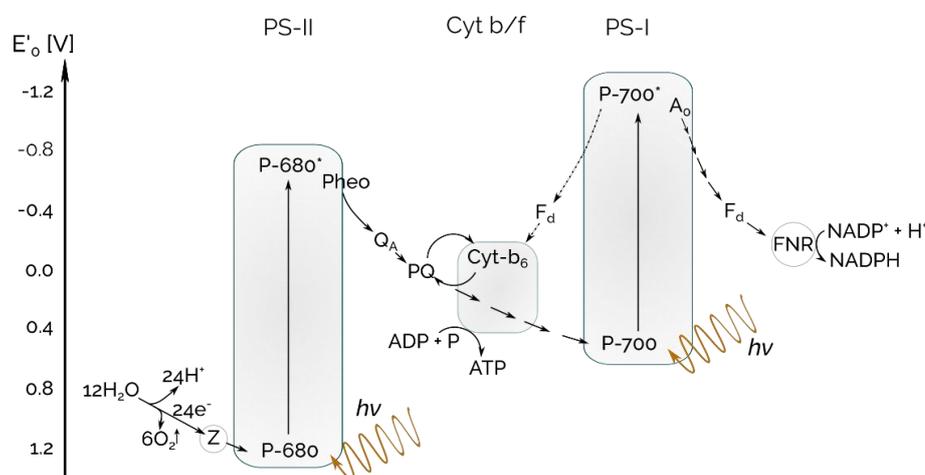
Figure 2-4: Simplified Z-scheme of the electron transfer chain performed during the light reaction in order to transmit electrons of the reductant to NADP$^+$ as a non-cyclic transfer pathway. In addition, a proton gradient across the thylakoid membrane leads to the photophosphorylation of ADP in order to generate ATP. The endogenic light reaction is catalyzed by the two photosystems after light absorption and conversion of the solar energy into chemical energy. The redox potential of the key compounds is indicated on the left. The scheme is reproduced based on the publication of Govindjee et al.[41]

The primary function of the light reaction is the oxidation of water and the synthesis of the reduction equivalents by transmitting protons and electrons to the coenzyme NADP$^+$ via complex reaction

cascade. The whole reaction pathway, known as the non-cyclic electron transfer pathway or Z-scheme, is characterized by the interaction of three different protein complexes – the photosystems I and II and the cytochrome complex as it is indicated in Figure 2-4. The light reaction is initiated by the oxidation of water, when the photosystem II withdraws the electrons from the reductant and oxygen is evolved in the interior of the thylakoid membrane. The resorbed electrons are first transmitted to pheophytin (*Pheo*), the first electron carrier intermediate, and subsequently to plastoquinone molecules (Q_A and *PQ*), which can move freely within the membrane. Thus, the electrons are transferred to the cytochrome complex (*Cyt-b_6*) and sequentially to the photosystem I. The second photosystem maintains the electron transfer pathway after light absorption. Finally, the electrons are transferred to the ferredoxin-NADP$^+$ reductase (*FNR*) that catalyzes the reduction of NADP$^+$. In addition to the electron transfer, a proton gradient is established between the lumen, the inner side of the thylakoid, and its surrounding, called stroma, when protons are accumulated in the lumen. This proton gradient is used by the ATP synthase for the photophosphorylation of ADP to generate ATP.[39,41,42]

As it is shown in Figure 2-4, the overall process of the light reaction is an endogenic process meaning that energy is required to catalyze the electron transfer pathway. This energy is provided by the solar energy which is absorbed by the photosystems I and II and sequentially converted into chemical energy that can be used in the electron transfer pathway. Thus, these photosystems act as light converter and electron pump across the thylakoid membrane and play therefore an important role within the photosynthetic process. The exact molecular structure of these photosystems is beyond the scope of this thesis. However, in general, the photosystems are composed of two main complexes – the light-harvesting complex, also called antenna complex, and the reaction center at the bottom side, which are linked to each other. Within the antenna complex, various photopigments are attached to membrane-embedded proteins and tightly packed to so called 'pigment collectives'.[39] These photopigments are auxiliary pigments, such as carotenoids, xanthophylls, other types of chlorophyll and phycobiliproteins, all featuring an extended conjugated π-electron system which enables the absorption of light at certain wavelength ranges. The main objective of the antenna complex is to collect solar energy, convert it into chemical energy and subsequently transfer the energy to the chlorophyll-dimer residing at the reaction center. This chlorophyll-dimer, known as 'special pair', catalyzes the photolysis of water and thus maintains the light reaction of the photosynthesis. The overall transmission process of the light energy from the antenna complex to the 'special pair' in the reaction center is called *energy trapping*. Figure 2-5A illustrates a simplified model of a photosystem and demonstrates its central role within the light reaction. All higher plants and algae exhibit only two photosystems, whereas cyanobacteria, red algae and glaucophytes have an additional antenna complex attached to the thylakoid membrane. This so called phycobilisome has a different structure compared to the photosystems I and II as it is shown in Figure 2-5B. All light absorbed by this additional antenna complex is transferred to the reaction center of the photosystem II.[39] Since further pigments are used for light absorption, the amount of absorbed energy is increased and the overall spectrum, where light can be absorbed, is extended compared to other algae. Thus, these additional photopigments enable the habitation of

areas with low light conditions. An overview of the absorptive behavior of the photopigments is shown in Figure 2-6.
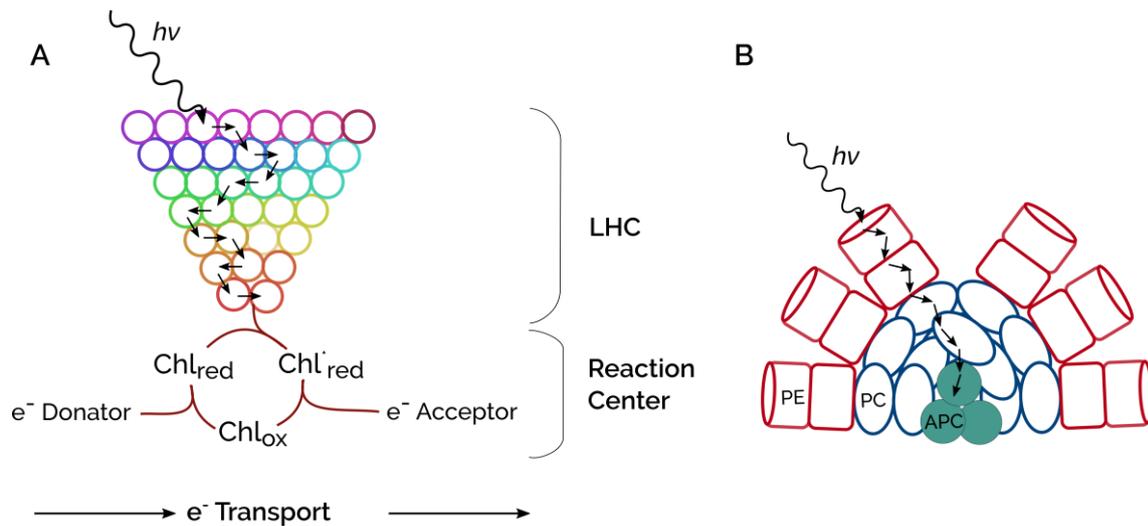
Figure 2-5: Simplified models of photosystems used in algae and cyanobacteria to maintain the electron transfer pathway. (**A**): General scheme of the photosystem I and II composed of the light-harvesting complex (LHC) and the reaction center. Auxiliary photopigments are tightly packed within the LHC enabling an enlarged absorption of the sunlight. The absorbed energy is subsequently transferred across an energy transport cascade to the chlorophyll-dimer in the reaction center, which maintains the electron pump of the light reaction. The scheme is reproduced from the publication of Schopfer et al.[39] (**B**): Scheme of the additional photosystem of cyanobacteria, red algae and glaucophytes which is attached to the thylakoid membrane. This peripheral photosystem is called phycobilisome and transfers energy absorbed by phycoerythrin (PE), phycocyanin (PC) and allophycocyanin (APC) to the reaction center of the photosystem II.
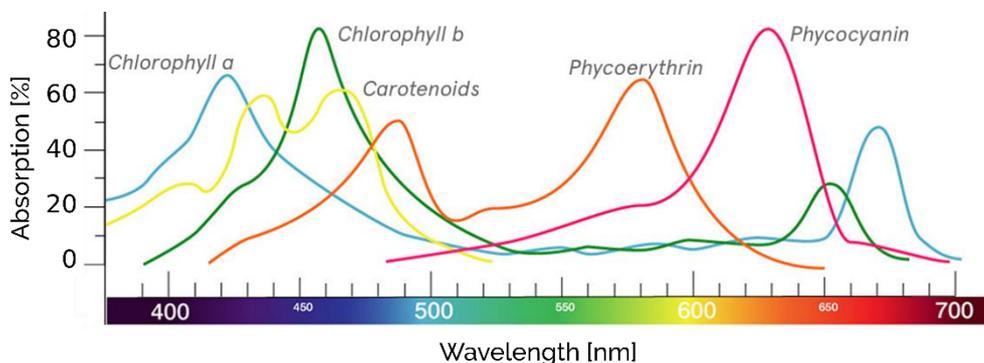
Figure 2-6: Absorption spectra of the major photopigments found in plants, algae and cyanobacteria. While chlorophyll-a is omnipresent in all photosynthetically active organisms, phycoerytrhin and phycocyanin are limited to cyanobacteria and some algae.

To sum up, accessory pigments are used to enlarge the spectral range from which energy is absorbed and therefore increase the efficiency of light absorption. Moreover, the photosystems are crucial to reduce the average reaction lifetime of photochemical reactions in a manner that the energy is made accessible for the biochemical reaction pathway. While biochemical reactions generally need few milliseconds or even seconds ($10^{-3} - 1\,\mathrm{s}$), the lifetime of photochemical reactions is much lower in the range of few femto- or picoseconds ($10^{-15} - 10^{-12}\,\mathrm{s}$).[39] In addition, these photopigments, in

particular carotenoids, have another benefit for the photosynthetic apparatus, as they protect the organisms from photodamage and photoinhibition due to reactive oxygen species that might occur under high light conditions. During a collection of non-photochemical quenching reactions, surplus energy is deactivated non-radiatively as heat.[43]

Although the class of pigments used in the photosystems remains the same and chlorophyll-a is the main photosynthetic pigment in all photoautotrophic organisms, the distinct pigmentation varies between different algal classes. Some photopigments are limited to certain algal groups and besides, the relative composition of the pigments in the light-harvesting complex adapts due to different environmental conditions, such as light conditions, nutrient availability and other stressors.[1,43] However, during the past decades, extensive studies enabled the identification of marker pigments that are characteristic for certain algal classes. Detailed information about these marker pigments and their absorption behavior is provided in section 2.3.2 and in chapter 5.

2.2.3. Energy Transfer and Trapping in Photosystems

Absorption is the initial step of photosynthesis. Photopigments within the light-harvesting complexes absorb part of the incident light and transfer the excitation energy to the reaction center. In general, excitation of a molecule or an atom is the promotion of one electron from the ground state to a higher energy level after absorption of energy. According to the Franck-Condon Principle, the electronic transition takes place while retaining the nuclear configuration of the atom or molecule. These electronic transitions, which are faster than any nuclear transformation, are called 'vertical transitions'. The energy required for excitation must coincide with the energy difference between the excited state (S_1 or S_2, marked as S^*) and the ground state (S_0) as it is defined by equation 2.1.

$$\Delta E = E_{S^*} - E_0 = h \cdot v = \frac{hc}{\lambda} \qquad \textbf{2.1}$$

In the photosystem II, the energy of red light is sufficient to stimulate the excitation from the ground state to the first excited energy level S_1, whereas the blue light stimulates the excitation to the second excited energy level S_2. However, the energy of the green light cannot be used for excitation in most photosystems, except in those photosystems containing phycobilisomes. In addition, the second excited state S_2 is less stable than the first excited state, wherefore the excited electron relaxes fast to the first excited state via internal conversion and emission of heat. Moreover, according to Kasha's rule, energy emission is only possible from the lowest excited state, wherefore the electron relaxes first to the lowest excited energy level by vibrational relaxation. Due to this non-radiative relaxation and deactivation of part of the excitation energy, the amount of surplus energy is decreased. Consequently, the emission is shifted to higher wavelengths compared to the excitation, which is known as Stokes shift. Starting from this lowest excited state, the electron tries to dispose further energy in order to regain its favored ground state S_0. Surplus energy can be emitted

subsequently through several non-radiative and radiative deactivation processes following the fundamental principles of thermodynamics summarized by Valeur[44] and Lakowicz[45]:

- o   Non-radiative transition by internal conversion and emission of heat k_{IC}
- o   Fluorescence emission k_f
- o   Intersystem crossing and phosphorescence k_{ISC}
- o   Photochemical reaction k_{PR} or energy transfer to a neighboring molecule k_{ET}

An overview of these electronic processes and a correlation between the electronic transients and the absorption and emission spectrum of chlorophyll-a is exemplarily shown in Figure 2-7.
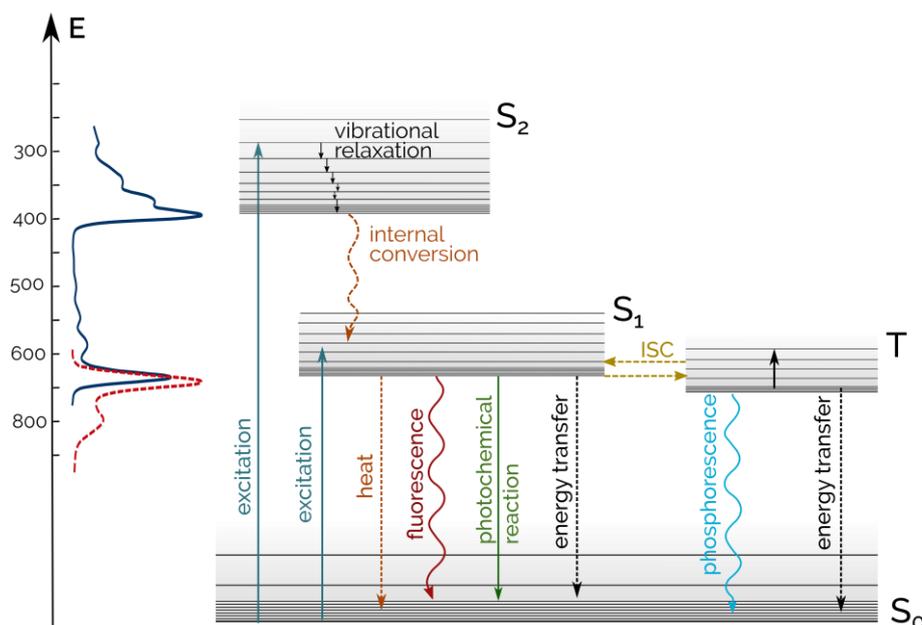
Figure 2-7: Jablonski diagram illustrating the electronic transitions within an atom or molecule after absorption of energy. Each energy level S_i is linked to several vibrational energy states, which are represented as horizontal lines. During excitation, one electron is promoted from the ground state S_0 to a vibrational level of one of its excited states (S_1 or S_2). Following the fundamental principles of thermodynamics, the excited electron tries to dispose surplus energy via diverse deactivation processes to regain its favored minimal energy level S_0. On the left, a typical excitation and emission spectrum of chlorophyll-a, marked as a blue, solid line and as a red, dotted line, respectively, links the electronic transitions and the spectrum.

Upon excitation and conversion to the ground state of the first excited energy level S_1, the relaxation processes mentioned above compete with each other as potential deactivation process in order to regain the favorable energy minimum of the atom or molecule. Thus, the lifetime of the excited state can be described by these deactivation processes according to equation 2.2. The corresponding fluorescence quantum yield Θ_f is then described relatively to the sum of all deactivation processes taking place (see equation 2.3).

$$\tau_{total} = \frac{1}{\kappa_{total}} = \frac{1}{\kappa_{IC} + \kappa_f + \kappa_{ISC} + \kappa_{PR} + \kappa_{ET}} \qquad \textbf{2.2}$$

$$\Theta_f = \frac{\kappa_f}{\kappa_{total}} = \frac{\kappa_f}{\kappa_{IC} + \kappa_f + \kappa_{ISC} + \kappa_{PR} + \kappa_{ET}} \qquad \textbf{2.3}$$

In solution, the *in-vitro* fluorescence quantum yield of chlorophyll-*a* can be determined to $\Theta_f = 0.3$, whereas the fluorescence quantum yield in living cells decreases by a factor of ten to $\Theta_f \sim 0.03$.[46] Thus, the maximal fluorescence quantum yield is reduced due to the fact that other deactivation processes prevail over fluorescence emission. However, intersystem crossing from the singlet state to the triplet state and subsequent phosphorescence to the ground state does not occur in cells under normal conditions ($k_{ISC} = 0$), as phosphorescence might encourage the formation of reactive oxygen species that can lead to photodamage and photoinhibition. Therefore, in order to protect the integrity of the photosystem, the triplet state of chlorophyll is avoided by certain quenching reactions.[39,41] While intersystem crossing is avoided under normal conditions, energy transfer between photopigments across the light-harvesting complex and charge separation within the reaction center are the dominant reactions in the photosystem to maintain the photosynthesis.

Energy transfer reactions are the prominent processes for photopigments within the light-harvesting complex to migrate the excitation energy to the reaction center.[47,48] Although the mechanisms for energy transmission are not fully clarified in their complexity, it is assumed that the Förster resonance energy transfer (FRET) is the prevailing mechanism within the light-harvesting complex. Förster proposed this energy transfer as non-radiative process between an excited donor molecule and an acceptor in the ground state as it is shown in Figure 2-8. During energy transition, the donor relaxes radiationless to the ground state, while the acceptor is excited. The efficiency of the FRET depends on spatial, physical as well as on spectral conditions: First, both molecules must be in close proximity, so that weak dipole-dipole interactions are possible over which the energy can be transmitted.[45] Second, the relative orientation of the dipole moment of the donor and the acceptor must be parallel during the singlet-singlet transfer, when the spins of donor and acceptor remain unchanged during transmission. The last condition for energy transmission is that the emission spectrum of the donor and the excitation spectrum of the acceptor sufficiently overlap in that manner that the acceptor is able to absorb a part of the emitted energy according to equation 2.1.
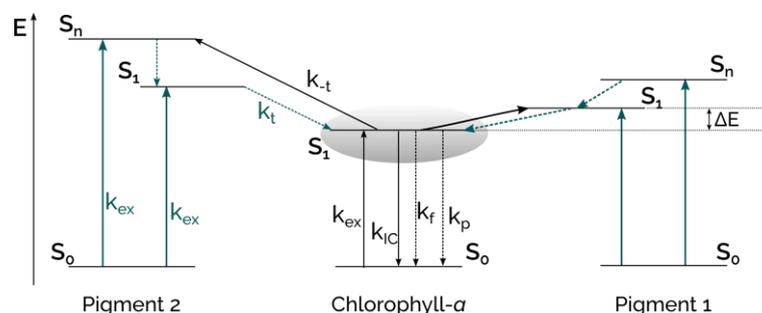
Figure 2-8: Jablonski diagram accounting the Förster resonance energy transfer (FRET) as prevailing mechanism to describe the energy migration across the light-harvesting complex to the reaction center. Upon excitation of the donor pigment the excitation energy is transferred from the donor to the acceptor via a non-radiative reaction and conservation of the total spin. In this case, the donor pigments are labeled as pigment 1 and 2, whereas the acceptor pigment is chlorophyll-*a*. The excited acceptor has several possibilities for the relaxation to the ground state as it was described previously.

The greater the energy difference between the excited states of the donor and the acceptor molecule, the more likely it is that the energy transition is a directed and irreversible transition.

Besides the Förster mechanism, Dexter proposed another mechanism that accounts for some inconsistencies, such as ultrafast transfer steps (about 200 fs) and the sometimes small overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor.[48,49] This Dexter mechanism bases on electrons hopping between the donor and the acceptor, wherefore the distance between these molecules must be even smaller than for the FRET. Moreover, a distinct overlap of the wave functions of the donor and the acceptor is crucial instead of a spectral overlap as it is the case for FRET.

After successful transmission of all excitation energy to the reaction center, the energy is trapped by the special pair as it is described above. Upon excitation, the chlorophyll-dimer has several possibilities for its relaxation to the ground state, as it is shown in Figure 2-7. However, as mentioned above, the fluorescence quantum yield in living phytoplankton is low with $\Theta_f \sim 0.03$. The reason for the small quantum yield is that, assuming that the photosystems are intact, fluorescence can only occur from the chlorophyll-dimer in the reaction center, since all excitation energy is effectively transmitted to the reaction center. By contrast, under normal conditions when the photosystem is intact, the photochemical reaction and charge separation prevails. Charge separation takes place, when the chlorophyll-dimer uses the trapped energy to withdraw electrons from the reductant in order to drive the synthesis of ATP and NADPH across the electron transfer chain and to maintain the photosynthesis.

To summarize, under normal conditions and when the photosystems are intact, these complex and densely packed photosystems enable algae and cyanobacteria to absorb energy from a broad wavelength range and consequently to inhabit areas of different light conditions all over the world. Moreover, the photosystems enable algae and cyanobacteria to adapt to different stressors such as changing light conditions or nutrient deficiencies.[10,43,50–52] One important method to study the competing relaxation processes of fluorescence and photosynthesis activity as well as the influence of different stressors is the time-resolved analysis of the chlorophyll fluorescence. The distinct shape of the chlorophyll fluorescence curve of dark adapted cells is characteristic and subject to comprehensive research projects. Although it is beyond the scope of this thesis, the chlorophyll induction kinetics has to be highlighted as an important research topic to consider for appropriate biomass quantification. For detailed information about the chlorophyll kinetics and the Kautsky effect, reference is made to Lazár, Stirbet and Govindjee.[53–55]

2.3. Pigment-based Chemotaxonomy

Chemotaxonomy is a biological classification method where algae are categorized according to similarities in the structure and the composition of distinct biochemical compounds. The pigment-based chemotaxonomic approach assumes that several photosynthetic pigments in algae are restricted to only one or few taxa. These pigments are called *'marker'* or *'diagnostic'* pigments which can be used for the identification and the quantification of bloom compositions after mathematical evaluation.[56–58] In addition, the applicability of a pigment-based identification of algae populations has been increased due to the development of high-performance liquid chromatography (HPLC) techniques and spectrofluorometric methods. Due to these reasons, the use of pigment data to study algae populations in the water columns has become more popular over the past decades.[59,60]

Simultaneously, different chemometric methods were developed to quantify and identify algal blooms and their composition in marine environments *in vivo* as well as *in vitro*.[59,61–66] These chemometric studies are based on a multivariate concept meaning that multiple information is combined and evaluated together. A comprehensive overview of different approaches used in literature, their characteristics and their challenges is provided by Roy et al..[64] These strategies enhance the information acquired from algal samples and have been proven to be a valuable tool for pigment-based identification of algae populations in the past. The main advantage of these methods is that more pigments can be analyzed, even when they overlap. As these photopigments are representatives for specific phytoplankton groups, the classification and discrimination is limited to the class level of algae.[56,57,59]

For our approach of a miniaturized and low-cost algae monitoring device, we focused on spectrofluorometric techniques combining several excitation wavelengths in the visible wavelength range. The device's main advantage is its convenient handling once the evaluation strategy and the required reference database are established. One challenge in developing the device was to identify relevant pigments that are reliable for distinct discrimination of cyanobacteria and dinoflagellates in mixed assemblages and to establish a comprehensive database of reference algal samples.

2.3.1. Feature Extraction

Initially, the spectral features of the algal classes were characterized according to their emission and excitation behavior in order to extract relevant regions that are most representative for certain algal groups. These differences are then used for multivariate discrimination of algae with special focus on the reliable identification of cyanobacteria and dinoflagellates.

In Figure 2-9, the principal excitation/emission behavior of cyanobacteria in contrast to other algae is shown in a fluorescence excitation-emission matrix. As it can be seen, the emission behavior of both, algae and cyanobacteria, is similar, whereas their excitation behavior differs significantly.

Algae as well as cyanobacteria emit mainly in the near infrared range between $660 - 720$ nm. While cyanobacteria can be excited mainly in the visible to near-infrared range between $550 - 670$ nm, the excitation spectrum of algae is extended to shorter wavelengths ($250 - 550$ nm).
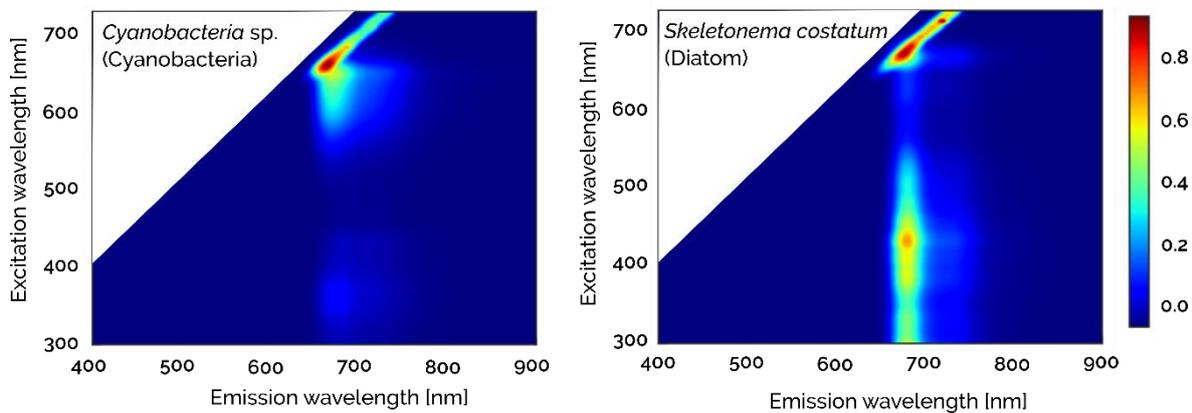
Figure 2-9: Fluorescence excitation-emission matrix of a cyanobacteria (*Cyanobacteria* sp.) and an algae (*Skeletonema costatum*) as representatives to study principal spectral differences between cyanobacteria and algae. The spectra were recorded on a Horiba fluorescence spectrofluorometer at room temperature. To avoid sedimentation during measurement, the dark-adapted samples were held in suspension with glycerin.

Thus, for an initial differentiation between cyanobacteria and algae, it might be appropriate to study the relative excitation intensity at shorter wavelengths compared to the chlorophyll excitation in the near-infrared range. For further separation of other algal groups, however, we investigated and compared the relative emission and excitation spectra of different groups.

2.3.1.1. Emission Spectra

The fluorescence emission of intact cells is observed from the reaction center, where chlorophyll-*a,* as ubiquitous main photosynthetic pigment, deactivates surplus energy from the lowest excited state.[1] Thus, the shape of the fluorescence emission spectra is quite similar for all algae and cyanobacteria, varying only slightly according to the distinct environment of the reaction center and depending on the extent of their state-transition (compare Figure 2-10).[67] However, the emission spectrum of cyanobacteria is slightly shifted to shorter wavelengths (hypsochromic shift), as additional fluorescence can be emitted from phycocyanin, which is part of the phycobiliproteins attached to the thylakoids of algae.[68] For further details on the pigmentation of the photosystems, please refer to section 2.3.2.
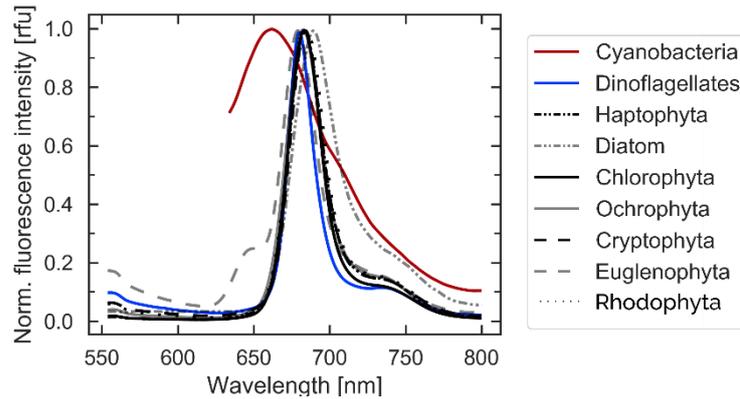
Figure 2-10: Overlay of the normalized fluorescence emission spectra of eight different phytoplankton groups. The spectra were recorded with a HITACHI fluorometer between 550 – 800 nm after excitation at 440 nm (algae) or 595 nm (cyanobacteria). The dark-adapted samples were measured at room temperature and they were held in suspension with glycerin to avoid sedimentation of the cells. After measurement, the spectra were normalized in each case to their maximal fluorescence emission peak.

2.3.1.2. Excitation Spectra

As initial fluorescence excitation-emission matrices suggested (see Figure 2-9), the excitation characteristics of cyanobacteria differ significantly from other algal groups. The photosystems of cyanobacteria are dominated by phycobiliproteins, which are blue or red chromophores attached to the thylakoids of their chloroplasts. These phycobiliproteins absorb mainly between 565 nm (phycoerythrin) and 650 nm (allophycocyanin).[1,69] Though rhodophyta contain phycobiliproteins as well, their relative pigment composition and, consequently, their relative excitation characteristics differ from those of cyanobacteria, especially in the range between 500 − 580 nm as shown in Figure 2-11.
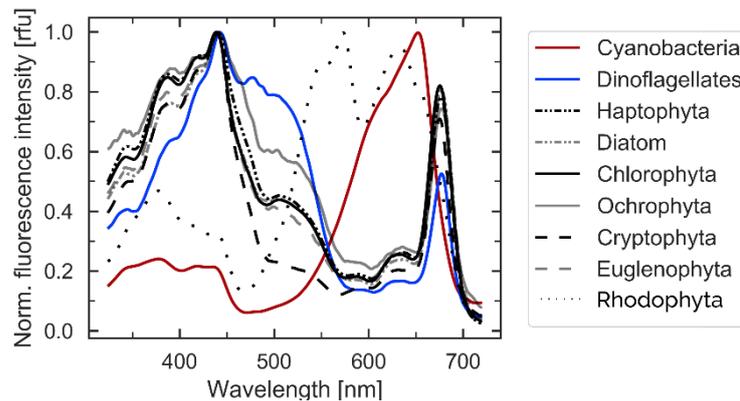
Figure 2-11: Overlay of the normalized fluorescence excitation spectra of eight different phytoplankton groups. The spectra were excited with a HITACHI fluorometer between 300 – 750 nm and their fluorescence emission was recorded at 750 nm at room temperature. The dark-adapted samples were held in suspension with glycerin to avoid sedimentation of the cells. After measurement, the spectra were normalized in each case to their maximal fluorescence excitation peak. The detector was covered with an RG-665 longpass filter to avoid spectral artifacts.

The average excitation spectrum of dinoflagellates differs slightly from other algae, especially in the wavelength range between 450 – 550 nm. In this wavelength range, dinoflagellates absorb more energy compared to other algae and exhibit therefore a broader absorption shoulder.

To summarize, while the emission spectra of algae and cyanobacteria are quite similar, the excitation spectra differ significantly in the wavelength range between 450 – 600 nm. Thus, it seems promising to discriminate algal groups rather based on their excitation characteristics than on their emission characteristics. Taking these considerations into account, we further extracted spectral key features that might be appropriate for reliable differentiation of cyanobacteria and dinoflagellates from other algae even in algal blooms of mixed composition. In the next step, these spectral key features were linked to diagnostic pigments.

2.3.2. Trapping Characteristic Marker Pigments

Although the relative pigment composition of microalgae is strongly dependent on several environmental factors, few pigments are known as unambiguous markers for certain algal groups.[70] These marker or diagnostic pigments are distributed across several taxa and they are characteristic for one or few algal classes. In few cases, diagnostic pigments might even be markers for certain algal species such as 4-keto-myxoxanthophyll and its ester which are diagnostic for the cyanobacterium *Nodularia spumigena*.[56,71]

In order to link the spectral characteristics investigated by spectroscopic analysis, that were presented in the previous section 2.3.1, with known diagnostic pigments, we analyzed the relative pigment composition of 53 different unialgal cultures belonging to eight different algal phyla. This study was done by HPLC analysis. An overview of these diagnostic pigments is given in Table 2-1. This table is completed by literature research as some water-soluble marker pigments, such as phycobiliproteins, cannot be found on the HPLC column. For further details about diagnostic pigments of different subphyla of dinoflagellates, please refer to chapter 5.

However, the overall absorption spectrum of an algal sample is not just the summation of individual absorption coefficients of all pigments present in their photosystems and pigment packaging effects, the cell size and pigment-protein interactions have to be considered.[23,66,72] Besides, variations in the pigment composition can occur due to varying light conditions, nutrient availability and growing conditions.[10,50] These effects influence the effective *in vivo* absorption of photosynthetic pigments leading to a wavelength shift of even 40 nm in case of fucoxanthin.[23] Thus, it is crucial to consider these influencing factors for the development of an appropriate *in-situ* monitoring device and a classification strategy based on spectral characteristics. Bidigare[23] and Bricaud[72] attempted to reconstruct such *in vivo* absorption spectra. They combined individual 'unpacked' absorption coefficients of pure pigments linearly and shifted these theoretical spectra until they matched with *in vivo* spectra. In Figure 2-12A, an example of this attempt is presented

by the diatom *Conticriba weissflogii*. Theoretical absorption coefficients of major pigments are plotted together with the excitation spectra of the unialgal solution. The accumulated absorption spectrum based on the theoretical absorption coefficients matches well with the measured excitation spectrum.

Table 2-1 Diagnostic pigments identified by HPLC analysis and completed by literature research. [a]

phylum	per	19HF	19BF	allo	dino	fuco	lut	zea	vau	phyco	chl-*b*	chl-*a*
absorption peak (nm)	475	448, 472	447, 471	452, 482	417, 441, 470	450	446, 475	429	419, 444, 471	492, 547	461, 597, 647	431, 617, 662
cyanobacteria	--	--	--	--	--	--	--	x	--	x[69]	--	x
rhodophytes[73]	--	--	--	--	--	--	x	x	--	--	--	x
dinoflagellates	x	x	x	--	x	x	x	x	--	--	x	x
diatom	--	--	--	--	--	x	--	--	--	--	--	x
haptophytes	--	x	x	--	--	x	--	--	--	--	--	x
ochrophytes	--	--	--	--	--	x	--	x	x[74]	--	--	x
chlorophytes	--	--	--	--	--	--	x	x	--	--	x	x
euglenophytes	--	--	--	--	--	--	--	--	--	--	x	x
cryptophytes	--	--	--	x	--	--	--	--	--	x[64]	--	x

[a] Visible absorption properties of pigments in eluent after HPLC analyses.[75–78] Abbreviations: ***per:*** peridinin, ***19HF:*** 19'-hexanoyloxy-fucoxanthin*, **19BF:*** 19'-butanoyloxy-fucoxanthin, ***allo***: alloxanthin, ***dino***: dinoxanthin, ***fuco:*** fucoxanthin, ***lut:*** lutein, ***zea:*** zeaxanthin, ***vau:*** vaucheriaxanthin, ***phyco***: phycobiliproteins including phycoerythrin, phycocyanin and allophycocyanin, ***chl-b:*** chlorophyll-*b*, ***chl-a:*** chlorophyll-*a*

Based on the distinct diagnostic pigments for nine phytoplankton phyla and corresponding *in-vivo* absorption characteristics, eight LEDs were selected for the ALPACA in the visible to near-infrared wavelength range: 380 nm, 403 nm, 438 nm, 453 nm, 472 nm, 526 nm, 593 nm and 640 nm. These LEDs are considered to be most significant for reliable identification of cyanobacteria and dinoflagellates. For further details on the emission characteristics of these LEDs, please refer to Table 4-1 in section 4.4.3. In Figure 2-12B, the normalized fluorescence spectrum of the diatom *Conticriba weissflogii* measured on the spectrofluorometer is compared to the discrete fluorescence pattern that is obtained from the ALPACA upon sequential excitation using these LEDs. The discrete fluorescence pattern matches well with the continuous excitation spectrum, although few LEDs are overestimated compared to others. This might be a problem of scattering effects in the ALPACA which is not sufficiently corrected but cannot be avoided completely in the setup. However, a correction of these artifacts is done during data evaluation by baseline correction. Besides, measurement uncertainties might occur due to varying pigmentations during growth as the algal sample was not measured consecutively on both instruments.
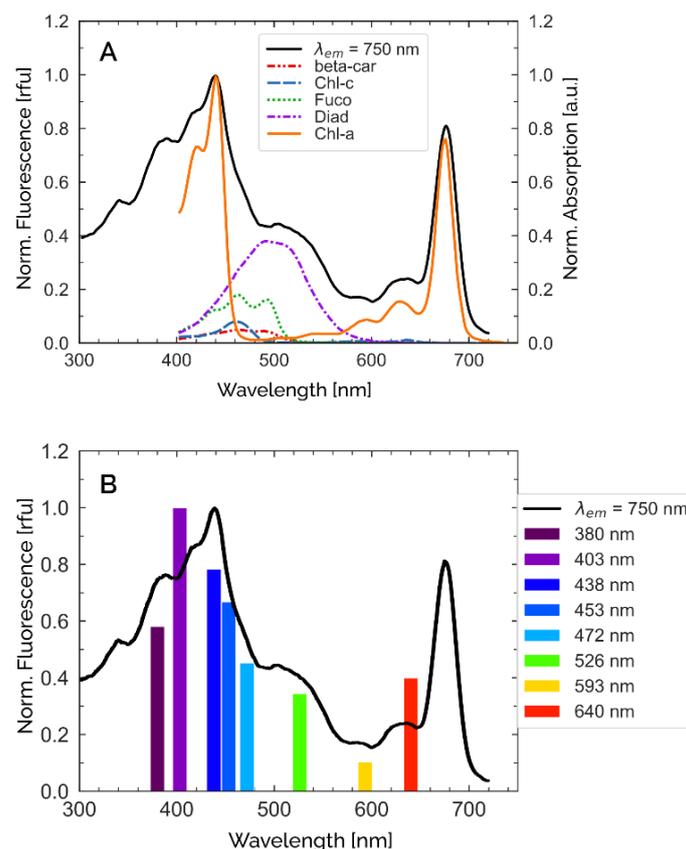
Figure 2-12: (**A**) Exemplary comparison of the normalized excitation spectrum with theoretical absorption coefficients of major diagnostic pigments to allocate spectral characteristics of the diatom *Conticriba weissflogii* based on Babin and Bidigare.[23,79] The fluorescence excitation spectrum of the dark-adapted unialgal sample was recorded with a HITACHI fluorometer at room temperature. The algal sample was held in suspension to avoid sinking of the cells. The emission wavelength was set to 750 nm. (**B**) Comparison of the normalized fluorescence spectrum and the discrete fluorescence pattern obtained from the ALPACA upon sequential excitation with eight different LEDs.

2.3.3. Mathematical Description

To discriminate algal groups based on their spectral features and to identify whether cyanobacteria or dinoflagellates are present in the sample, a mathematical description is required. Further, the chosen method must handle a large amount of data, while each sample data is characterized by a set of measurement features. In addition, graphical representation is often required for data interpretation but impossible if more than four measurement features are examined. Thus, these specific questions yield into the field of multivariate data analysis or rather into the field of discriminant analysis.[80–82]

2.3.3.1. Theory, Assumptions and Challenges

For our purpose, the analysis of choice was Fisher's linear discriminant analysis (LDA) – a statistical, supervised pattern recognition method. This method targets the optimal discrimination

of objects belonging to different groups and subsequently the appropriate allocation of an unknown sample to one group. The LDA seeks linear combinations of several measured features in order to separate objects belonging to different groups. This separation is done by a coordinate transformation and projection onto a lower-dimensional (sub-)space. In this subspace, the separation of groups and the classification of an unknown sample data are based on Fisher's criterion. This criterion implies that the separation between two or more groups is optimal when the distribution between groups is maximal, while the distribution of objects belonging to the same group must be minimized.[83] Subsequently, the Mahalanobis distance between an unknown sample and the group median describes the spectral similarity and can further be transferred into a probability of group membership.[84]

As a supervised pattern recognition algorithm, the LDA requires a set of well-characterized training data to learn hidden patterns and relations of statistically independent training objects. Hence, one challenge is to map the wide variety of spectral features of all algal groups which requires a substantial knowledge about spectral characteristics of algae. The identification of representative training objects was therefore being done by experts of the French Research Institute for Exploitation of the Sea (IFREMER) as well as in cooperation with the Plant biology and Ecology Department of the University of the Basque Country (UPV/EHU) in Spain.

In addition to the establishment of a well-characterized training database, it is assumed that the measurement features are normally distributed and statistically independent. However, this is not the case in real datasets, as normally data correlate slightly with each other. A measure of the dimension of this correlation is provided by the variance-covariance matrix or its standardized form, the correlation matrix. The entries of the variance-covariance matrix describe the statistical dependency of two measurement variables X_i, X_j (equation 2.4).

$$Cov(X_i, X_j) = \begin{pmatrix} Cov(X_1, X_1) & \cdots & Cov(X_1, X_n) \\ \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & \cdots & Cov(X_n, X_n) \end{pmatrix} \qquad \textbf{2.4}$$

with $Cov(X_i, X_j) = Cov(X_j, X_i)$ and $Cov(X_i, X_i) = Var(X_i)$

The dimension of the variance-covariance matrix n corresponds with the number of measurement features and is defined by the number of excitation channels in the ALPACA ($n = 8$). For the discriminant analysis, it is further assumed that each algal group has identical variance-covariance matrices.[82] An example of the variance-covariance matrix calculated for the training objects is given below in Table 2-2. The statistical dependency can be analyzed column by column by comparing all values with each other. For example, in the first column of Table 2-2, LED-380 is more dependent on LED-403 than it is self-dependent, whereas LED-403 (second column) is independent of other LEDs. This can be explained by the general setup of the photosystems. LED-403 targets the excitation of the Soret-band of chlorophyll-a and other LEDs are selected to excite further minor or major marker pigments. However, the blue wavelength range is dominated by the chlorophyll absorption as chlorophyll-a is the major photosynthetic pigment in the photosystems of algae and cyanobacteria. Therefore, LEDs emitting in the blue wavelength range might excite chlorophyll to

some extent besides other marker pigments. For detailed information about operational features for the LDA, please refer to the supporting information of chapter 5.

Table 2-2 Variance-covariance matrix across all 52 training objects indicating whether two LEDs are statistically dependent. The higher the (absolute) value, the higher the dependency.

	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
380 nm	0.031							
403 nm	0.054[a]	0.221						
438 nm	0.008	0.024	0.188					
453 nm	0.006	-0.037	0.090	0.190				
472 nm	-0.004	-0.086	0.007	0.113	0.195			
526 nm	-0.033	-0.095	-0.198[a]	-0.209[a]	-0.033	0.578		
593 nm	-0.038	-0.092	-0.051	-0.034	-0.037	0.060	0.101	
640 nm	-0.024	0.010	-0.069	-0.120	-0.153[a]	-0.070	0.092	0.334

[a] statistical dependent

Using LDA for the chemotaxonomic discrimination of algae based on their spectral features, there are a few challenges to consider. One challenge is the wide variety of spectral features within one algal group and various phenomena that affect the characteristic pigmentation of the individual algal sample, such as endosymbiosis or kleptoplasty. Kleptoplasty is a symbiotic phenomenon when host organisms ingest algae and restrain their chloroplasts. These chloroplasts remain temporally functional and can be used for photosynthesis. Subsequently, the host organisms feature the same spectral properties as the algae and therefore might be considered as algae by mistake. It is known in literature that kleptoplasty can be conducted by dinoflagellates themselves, where the kleptoplastids remain photosynthetically active for several months. In this case, the dinoflagellate cannot be identified correctly.[85] Additionally, the relative pigment composition depends on various environmental factors such as light intensity and quality, day length and nutrient availability as well as on the growth phase of the algae.[10,64] These influencing factors are not yet included into the training database, however it might be possible after detailed and comprehensive analysis. Thus, although it is not possible to prevent every false-negative event, the training database can steadily be extended including an increasing number of spectral alterations.

Besides the direct alteration of the pigment pattern in algae, artifacts and other interferences from suspended particles, yellow substances or colored dissolved organic matters might influence the reliability of the algorithm. Therefore, the ALPACA is mounted with optical filters to attenuate light from interfering factors.

2.3.3.2. Data Preparation

Before conducting the data analysis, a pre-treatment is required to enhance inter-comparability within the ALPACA as well as the reproducibility of experiments. The implemented standardization and correction strategies aim to balance influencing factors of several optical compounds, such as the transmission behavior of the mounted filters or the absolute light intensity of the chosen LEDs. The absolute peak wavelength of the LEDs differs within one batch and the peak wavelength is temperature depending. In order to balance differences resulting from the fabrication, the LEDs must be calibrated against an internal quantum counter, whereas the temperature dependency is controlled on the device.

For the correction of the mounted emission filters, an aqueous solution of sulforhodamine 101 was used as fluorescence standard with a final concentration of 19.7 μM. The average fluorescence emission was measured at each channel upon excitation with an LED-472 and was then normalized. In Figure 2-13A, the fluorescence of sulforhodamine 101 is compared with the relative transmission of both longpass filters. In general, the fluorescence signal recorded on the longpass filter RG-665 is four times higher than on the longpass filter RG-9. These correction factors are used for subsequent measurements in order to balance channel-specific differences of the longpass emission filters.
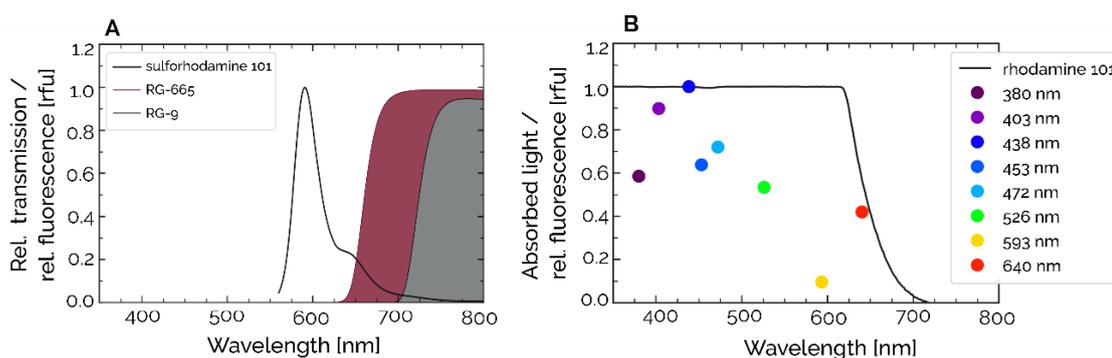
Figure 2-13: Correction of measured raw data is done on the emission and on the excitation site. (**A**) For the emission site, the fluorescence standard sulforhodamine 101 was recorded on each of the four silicon PIN photodiodes covered with two different longpass filters (RG-665, RG-9). For correction of the fluorescence emission, the transmission behavior of both longpass filters are taken into account. Normally, the fluorescence signal recorded on a detector covered with an RG-665 filter is four times higher than if the detector is covered with an RG-9 filter. (**B**) To balance differences in the emission behavior of the LEDs, rhodamine 101 (in ethylene glycol, 12 mM) was used as quantum counter. The signals recorded on the ALPACA are compared to the absorption spectrum measured on the photometer.

For the correction of spectral differences of the LEDs, correction factors were determined using rhodamine 101 (in ethylene glycol, 12 mM) as common quantum counter. The quantum yield of this standard must be 1, meaning that all incident light is re-emitted independently of the excitation wavelength.[44] For rhodamine 101, this is true in the wavelength range between $300 - 600$ nm as it is shown in Figure 2-13B. Thus, rhodamine can be used as quantum counter for all LEDs except for LED-640. However, in this case, the fluorophore operates as fluorescence standard. Therefore, the concentration of the standard must be defined correctly for correction purposes and must remain

constant for recalibration purposes. The fluorescence signal recorded on the ALPACA is compared to the amount of absorbed light $(I_0 - I)/I_0$ determined by the photometer according to equation 2.5. The ratio of the set value at certain wavelength and the actual value given upon excitation with the corresponding LED, yield into the correction factor of the LED κ_{LED} (compare equation 2.6).

$$I_{set}(\lambda) = \frac{I_0 - I}{I_0} = 1 - 10^{-A(\lambda)} \qquad \text{2.5}$$

$$\kappa_{LED} = \frac{I_{SET}(\lambda)}{I_{ACTUAL,\ LED}} \qquad \text{2.6}$$

where $A(\lambda)$ is the absorption of the standard at certain wavelength. $I_{set}(\lambda)$ is the set value or the amount of absorbed light at wavelength λ and $I_{actual,\ LED}$ is the emission recorded on the ALPACA upon excitation with the corresponding LED.

Then, the average fluorescence signal of the sample is determined and a baseline correction is conducted in order to reduce artifacts coming from biofouling grown in the capillary of the ALPACA. It is not possible to avoid biofouling completely, although multiple approaches to prevent biofouling in the system have been implemented, among others, cleaning steps in between experiences, periodic mechanical maintenance of the ALPACA and preservatives used as storage solution. To cope with this issue of the biofouling, we included a baseline correction in the data preparation procedure as it is shown in Figure 2-14. After the initial cleaning of the capillary, the sample is pumped through the capillary. In the transition between cleaning and sample solution, a sudden increase of the fluorescence signal is observed, when air is passed through. After a short conditioning phase, the fluorescence signal emitted from the biomass in the capillary remains constant over time. This time range is then used for evaluation.
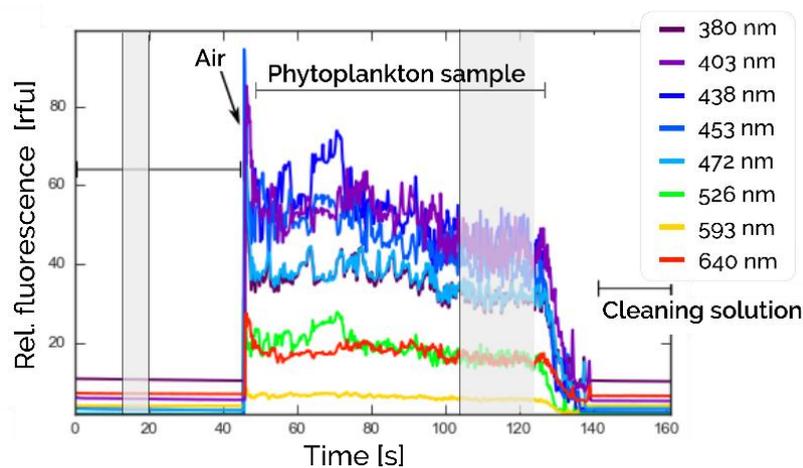
Figure 2-14: Initially, a cleaning solution is pumped through the capillary following by the algal sample, in this case *Skeletonema costatum* (Diatom). In the transient between cleaning and sample solution a sudden increase of the signal is observed when air is passed through. The time ranges marked in grey are used for evaluation and respectively for baseline correction in order to avoid artifacts coming from biofouling in the capillary.

Subsequently, the averaged and corrected fluorescence signal is normalized. The resulting bar plot, shown in Figure 2-15, matches the excitation spectrum recorded with a photometer as demonstrated before and coincides with the relative pigment composition of the algal sample.

The data preparation is done for all training objects and samples. In the following, the separation of training objects is done according to Fisher's criterion as described in the following section.
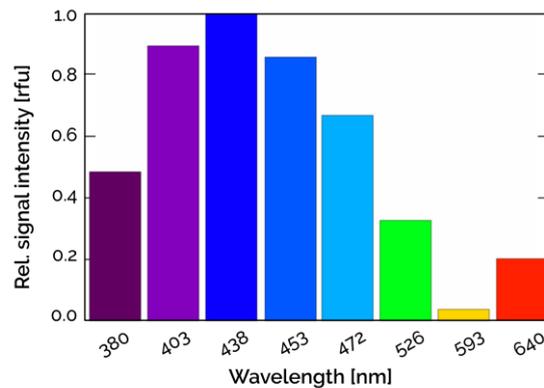
Figure 2-15: Resulting bar plot of the algal sample *Skeletonema costatum* (Diatom) representing the average fluorescence intensity upon excitation at eight different LEDs after pretreatment. The discrete bar plot displays the relative composition of marker pigments as discussed earlier.

2.3.3.3. Fisher's Criterion

After the data preparation, objects belonging to different groups must be separated according to Fisher's criterion. As mentioned before, Fishers criterion implies that the separation between two or more groups is optimal, when the scatter within one group is minimized, while in each case the scatter between two groups is maximal. Further, the separation is based on multiple measurement variables that are representative for these training objects aiming to emphasize significant differences between the groups.[83]

In Figure 2-16, the separation principle is described for a two classes problem for the sake of clarity, although it can be generalized for a multiple classes problem. Initially, we assume a set of n training objects $w = \{w_1 + \cdots + w_n\}$ where each object can be described as a weighted linear combination of m measurement variables φ_i:

$$w = \sum_{i=1}^{m} \alpha_i \cdot \varphi_i \qquad \textbf{2.7}$$

where α_i are the weighting factors. In case of the ALPACA, we used eight different LEDs as measurement variables and include 52 different training objects belonging to eight different algal classes for initial study. An overview of all training objects is given in the supporting information of chapter 5. Furthermore, as a supervised pattern recognition algorithm, the LDA requires information about the group membership of the training objects in order to seek discriminant functions that separate best objects belonging to different groups. For a multiple classes problem,

the LDA seeks multiple discriminant functions, in particular $K - 1$ functions, where K is the number of classes included into the training database. In our case, the LDA seeks seven discriminant functions, as eight different classes are included into the training database.
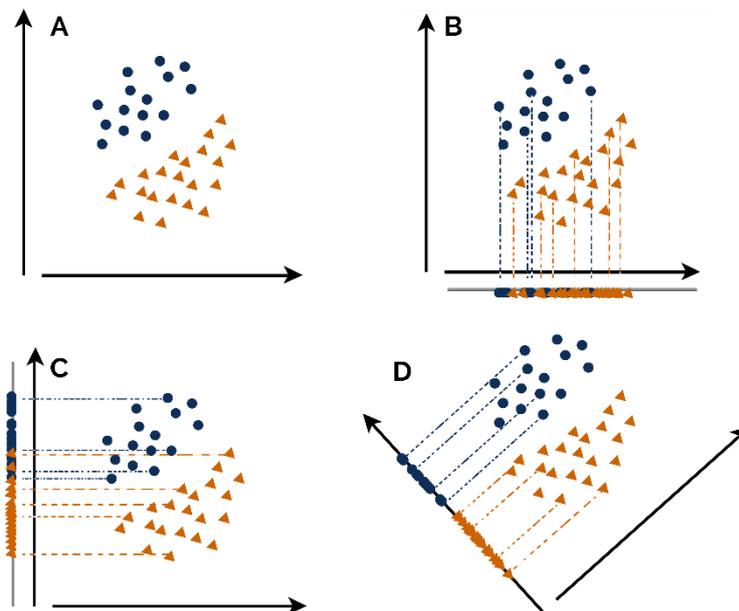
Figure 2-16: Mathematical principle of Fisher's linear discriminant analysis for a two classes problem. (**A**) In the original measurement space, the two datasets are close together. The projections onto one of the two original coordinate axes describe the distribution of the data for each variable and each group. As shown in the subfigures (**B**) and (**C**), there is a significant overlap of the two datasets. Thus, a complete discrimination is not possible in this case. However, the rotation of the coordinate system, also called coordinate transformation, enables the discrimination of the two sets after projection onto one axis (**D**). The optimal separation is calculated by the optimization and eigenvalue problem described in Fisher's criterion.

Following Fisher's criterion, the separation is optimal, when the within-scatter is minimal and the between-scatter is maximal. Thus, the overlap between two groups ought to be minimal (compare Figure 2-16C-D). This criterion can be transferred into a mathematical optimization problem as shown in equation 2.8:

$$J(w) = \frac{w^T \cdot S_B \cdot w}{w^T \cdot S_w \cdot w}$$

2.8

where S_W, S_B are the *within* and *between* scatter matrices of training objects belonging to different classes. To optimize the separation, a linear combination of the measurement variables φ_i must be determined, so that variables mapping significant differences of relevant algal classes must be emphasized compared to others. Therefore, optimal weighting factors must be determined for the linear combinations w. To solve the problem, the optimization problem of Fisher's criterion is transferred into an eigenvalue problem by derivation of equation 2.6:

$$\frac{\partial}{\partial w} J(w) = \frac{\partial}{\partial w} \frac{w^T \cdot S_B \cdot w}{w^T \cdot S_w \cdot w} = \frac{(w^T \cdot S_B \cdot w)' \cdot (w^T \cdot S_W \cdot w) - (w^T \cdot S_W \cdot w)' \cdot (w^T \cdot S_B \cdot w)}{(w^T \cdot S_W \cdot w)^2}$$

2.9

$$\frac{\partial}{\partial w} J(w) = \frac{2S_B \cdot w \cdot (w^T \cdot S_W \cdot w) - 2S_W \cdot w \cdot (w^T \cdot S_B \cdot w)}{(w^T \cdot S_W \cdot w)^2} = 0$$

The derivation becomes zero when its numerator becomes zero. Therefore, the denominator was neglected in the following and the whole equation was simplified by division of the factor 2.

$$\Rightarrow S_B \cdot w \cdot (w^T \cdot S_W \cdot w) - S_W \cdot w \cdot (w^T \cdot S_B \cdot w) = 0 \mid (w^T \cdot S_W \cdot w)^{-1}, \text{with } w, S_w \neq 0 \qquad \textbf{2.10}$$

Subsequently, equation 2.10 was divided by the initial denominator leading to further simplifications. Assuming that the within scatter matrix S_W is not a singular matrix and its inverse matrix exists, the optimization problem can be described as a known eigenvalue problem:

$$\Rightarrow S_B \cdot w - S_W \cdot w \cdot \lambda = 0 \Leftrightarrow (S_B - S_W \cdot \lambda) \cdot w = 0, \text{with } w, S_w \neq 0 \qquad \textbf{2.11}$$

The solution of the eigenvalue problem is given by its determinant $det(S_B - S_w \cdot \lambda)$. The eigenvalues and corresponding projection matrices span a characteristic space, into where the training objects are projected and separated best. As mentioned previously, there are a maximum of $K-1$ eigenvectors with corresponding eigenvalues. The relationship among different algal classes or centroids can be studied in this sub-dimensional score space. The first eigenvalues and corresponding projection axes in this space give the largest measure of Fisher's criterion, emphasizing the most prominent difference of the algal classes.

2.3.3.4. Group Allocation due to Mahalanobis Distance

After analyzing the pattern of the training objects, the group membership of an unknown sample can be predicted. For this purpose, the same projection is applied to the sample data as it was done for the training objects and the distance between the median of the group centroid and the sample score is calculated pairwise. The distance characterizes the spectral similarity between the sample and a particular algal group. The smaller the distance, the closer the spectral similarities. The distance is calculated using Mahalanobis distance as it is a scale invariant and translational balanced measure. Thus, the distance between the sample and each group remains the same, even after coordinate transformation and projection into the subspace. The Mahalanobis distance is calculated according to equation 2.12:

$$\partial(x, k_i) = \sqrt{(x - k_i)^T \Sigma^{-1}(x - k_i)} \qquad \textbf{2.12}$$

where x is the sample score, k_i is the median of class i and Σ^{-1} is the variance-covariance matrix which is calculated according to equation 2.1. The distance can further be converted into a probability of group membership assuming a Gaussian distribution curve of all data in each dimension and after auto-scaling. For auto-scaling, the distance is first mean centered and then divided by the variance.

$$\rho(\%) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma}} \cdot e^{-t} \quad \text{with } t = \frac{(\sigma - \mu)^2}{2\sigma^2} \qquad \textbf{2.13}$$

During the work for this doctoral thesis, this data evaluation strategy was implemented into a graphical user interface (GUI). A brief manual for this GUI is presented in chapter 6 and a validation of the mathematical principle using LDA for group allocation of unknown algal samples is published in a peer-reviewed journal and presented in chapter 5. Furthermore, one real-world application was conducted in the frame of the SCHeMA-project between February 14 – March 13 2017. In the port of Genoa, the long-term stability of the system was tested and a continuous identification of algae populations was conducted in the harbor. These results were published in a peer-reviewed journal (compare chapter 4). Further results gained from field trips in France are presented in chapter 9.

3. Dual Optical Sensors and Estimation of the Error

This chapter introduces the necessary theoretical background of the (modified) dual lifetime referencing (DLR) method based on a multifrequency phase modulation for simultaneous determination of two different analytes. It further provides a brief overview of challenges and advantages of those dual optical sensors and finally applies the DLR technique to four different hybrid sensing projects. In addition, this chapter provides an introduction for the estimation of the maximal error due to systematic and random bias that might occur during the measurement and their propagation across the evaluation.

3.1. Advantages of Hybrid Sensors

Optical sensors have proven to be valuable for elucidating and enhancing our understanding of dynamic, biological processes in (marine) life. Over the past decades, the demand for adaptive multiparametric sensing has increased in marine science and biotechnology, as the trend goes towards comprehensive sensor observation networks.[86,87]

In particular, optical sensors based on fluorometric methodologies are advantageous as fluorescence spectroscopy is a highly sensitive and selective methodology allowing the monitoring of the analyte of interest on a high-resolution scale in space and time.[86] Moreover, fluorescence based optical sensors meet the requirements of a good measurement performance in terms of reliability, robustness and versatility. Since light is the carrier of information, the sensory unit can be separated from the detector, allowing thus a contactless monitoring of the analyte of interest. As a result, this facilitates the development of versatile and adaptive sensor designs including planar sensors, sensor paints and dispersible nanosensors with real-time response.[88] This further enables the development of less sophisticated and low-cost instrumentations that are minimally invasive and feasible for miniaturization, when the indicator layer and the detector unit are decoupled. This versatility enables thus the application of these sensors in marine science, as they can be integrated into pressure stable and waterproof housings. According to literature, optical sensors have proven to be an attractive and efficient tool for measurement routines and real-time monitoring in marine science, medicine and biotechnology.[86]

The great advantage of optical sensors, in particular compared to electrochemical sensors, is their possibility of multiparametric sensing and multiplexing which is an elegant way to simultaneously determine multiple analytes within the same sensor. Since light is the signal carrier, it bundles multiple information, such as light intensity, luminescence decay time and polarization, and transmits the superimposed signal across the medium to the detector. Subsequently, the signal is

demultiplexed and demodulated in order to re-extract the relevant signal information of the indicator dyes.[88–90] The trend of multiparametric sensing takes into account that even single optical sensors must be corrected for diverse interfering factors, such as temperature or other external quenchers.[89]

The most widely used techniques in optical sensing base on the determination and monitoring of the absorbent and luminescent properties of the indicator, although the direct detection of luminescence intensity is often substituted by lifetime-based sensing methods.[87] According to Parker's law, the detection of the luminescence intensity depends on fluctuations in the sample properties, wherefore the detection of an intrinsically referenced parameter, such as the luminescence decay time is favored.[91,92] Moreover, optical (hybrid) sensors based on the detection of the decay time show a good performance in terms of sensor drift stability and longevity. Moreover, the investigation of the dual lifetime referencing technique applied in the frequency domain enables the determination of the decay time with less sophisticated instrumentations.

3.2. Challenges for Hybrid Sensors in Oceanography

Despite the fact that optical hybrid sensors are a promising tool for the simultaneous determination of multiple analytes, some requirements and challenges must be taken into account in the development of new optical hybrid sensors and sensor materials applied to marine science.

In oceanography, sensors are applied to a complex matrix consisting of several chemical and biological components, leading to high material standards and requirements concerning the sensor performance. Besides a high accuracy, the sensor must show an excellent precision and a low cross-sensitivity to other substances. Moreover, biofouling due to microbiological growth on the sensor surface must be avoided and cleaning strategies should be considered for long-term deployments without recalibration or maintenance. In addition, the sensor materials must show an outstanding drift stability throughout the deployment. While such sensors already exist for single optical sensors, the challenge for hybrid sensors is to combine those individual sensors or to develop new sensor materials with similar properties and overlapping spectral properties enabling the simultaneous excitation of the indicator dyes.[93–95]

Another challenge for multiplexed optical sensors is the development of appropriate evaluation strategies for signal processing and calculation of the analyte concentration with less calibration effort. Moreover, due to the iterative character of the evaluation strategies for multiplexed sensors, the error propagation and, in particular, the sensitivity and specificity of the second sensor are prone to errors. Thus, for an optimized multiparametric analysis, all available information that are transmitted by light, must be taken into account for the concrete hybrid sensor, including light intensity and luminescence decay time. As mentioned above, intensity based measurements are simple in terms of instrumentation, but their accuracy is often compromised as the fluorescence

intensity depends on various physicochemical properties of the sample and different instrumental conditions as it is defined by Parker's law described in equation 3.1.[92] Parameters that might affect the overall fluorescence intensity include the coloration, the turbidity or the refractive index of the sample as well as the geometry of the sensor and the absolute light intensity used for excitation.

$$I(\lambda_{em}) = \Theta(\lambda_{ex}) \cdot I(\lambda_{ex}) \cdot \varepsilon(\lambda_{ex}) \cdot \kappa \cdot l \cdot c \qquad \textbf{3.1}$$

where $I(\lambda_{ex}), I(\lambda_{em})$ are the luminescence intensity of the incoming and the emitted light, respectively. $\varepsilon(\lambda_{ex})$ is the molar absorption coefficient, $\Theta(\lambda_{ex})$ is the quantum yield, c is the effective concentration of the fluorophore and l, κ are the optical path and a proportional factor describing optical properties of the sample and the optical configuration of the instrument. This equation assumes an appropriate concentration of the fluorophore and the absence of inner filter effects or alterations of the sample properties due to agglomeration, photobleaching or leaching.

In terms of instrumentation, intensity based measurements are favored over lifetime measurements, in particular when the decay time of the hybrid sensor and thus the modulation frequency and the superimposed phase shift are small. Moreover, the measurements are prone to errors due to electromagnetic interferences, when the modulation frequency is in the range of the utility frequency. Consequently, appropriate strategies for standardization and internal calibration are crucial for reasonable analysis of multiparametric sensing and real-time monitoring, as well as a careful analysis of the error propagation across the evaluation.

3.3. Modified Dual Lifetime Referencing in the Frequency-Domain and Estimation of the Relative Error

As emphasized before, monitoring of intrinsically referenced parameters of two different indicator dyes incorporated into one dual optical sensor is highly attractive and an elegant method to apply. However, different approaches exist for the dual lifetime referencing technique operating either in the frequency domain or in the time domain. While this chapter focuses on the DLR technique operating in the frequency domain, reference is made to other publications describing the (modified) DLR technique operating in the time domain.[96,97]

Aside from the sophistication of the chosen methodology, the estimation of the maximal error is crucial in quantitative analysis, since errors always occur during quantification due to biases between the observed and the true value.[81,98] Moreover, errors can increase indeterminately due to the iterative character of various multivariate approaches.[82] Thus, it is significant to consider the error propagation across the evaluation and estimating the maximal error.

3.3.1. Principle of the (modified) Dual Lifetime Referencing Technique

The general principle of the dual lifetime referencing (DLR) technique was published by Klimant[99] and Wolfbeis et al.[86]. Borisov and Neurauter modified the intrinsically referenced measurement of the decay time for simultaneous determination of two luminescent indicators with overlapping spectral properties, but largely different decay times. For a description of the mathematical principles in more detail, references are made to their publications.[88,100]

In the frequency domain spectroscopy, the analyte-sensitive indicator dye is excited with a sinusoidally intensity-modulated light. The indicator re-emits the absorbed light at the same modulation frequency but with amended amplitude due to demodulation and with a delay in time. Thus, the response function of the indicator exhibits a phase shift compared to the excitation light that can be correlated to the decay time of the indicator and respectively to the analyte concentration. In case of a single-exponential decay this correlation is described by equation 3.3:

$$\tau = \frac{\tan(\varphi)}{2 \cdot \pi \cdot f_{\mathrm{mod}}} \qquad\qquad \textbf{3.3}$$

where τ is the decay time of the indicator sensitive to the analyte concentration, f_{mod} is the modulation frequency and φ the corresponding phase shift. For most analytes, the indicator dyes exhibit a short decay time of only a few nanoseconds (10^{-9} s). While a straightforward instrument can determine longer decay times, it fails in determining faster decay times. To precisely determine the decay time of the indicator, a more complicated, expensive and bulky equipment is required, which makes it unfavorable to apply. In order to overcome this challenge, the dual lifetime referencing technique was invented, combining the analyte-sensitive dye with a long-lived luminophore acting as reference dye. This reference dye exhibits a constant decay time of a few milliseconds and is insensitive to the analyte. Moreover, the reference dye and the indicator dye ought to have overlapping spectral properties enabling the simultaneous excitation of both dyes at the same modulation frequency. Therefore, the measured phase shift is an overlay combining the decay time of both luminophores, the constant decay time of the long-lived reference dye and the variable decay time of the short-lived indicator dye, with respect to their specific intensity ratio. The correlation between the overall phase shift Φ and the luminescence decay time of the long-lived reference dye and respectively the intensity ratio of the fluorescence indicator and the reference dye was published before by Huber et al. and is described by equation 3.2.[101] Considering equation 3.3, the decay time of the reference dye reveals a constant background signal with a constant shift in phase φ_P. Due to this correlation, the overall phase shift Φ is again in the range of a few micro- or even milliseconds enabling the usage of low-cost equipment feasible for miniaturization.

$$\cot(\Phi) = \cot(\varphi_P) + \frac{1}{\sin(\varphi_P)} \cdot \frac{A_F}{A_P} \qquad\qquad \textbf{3.2}$$
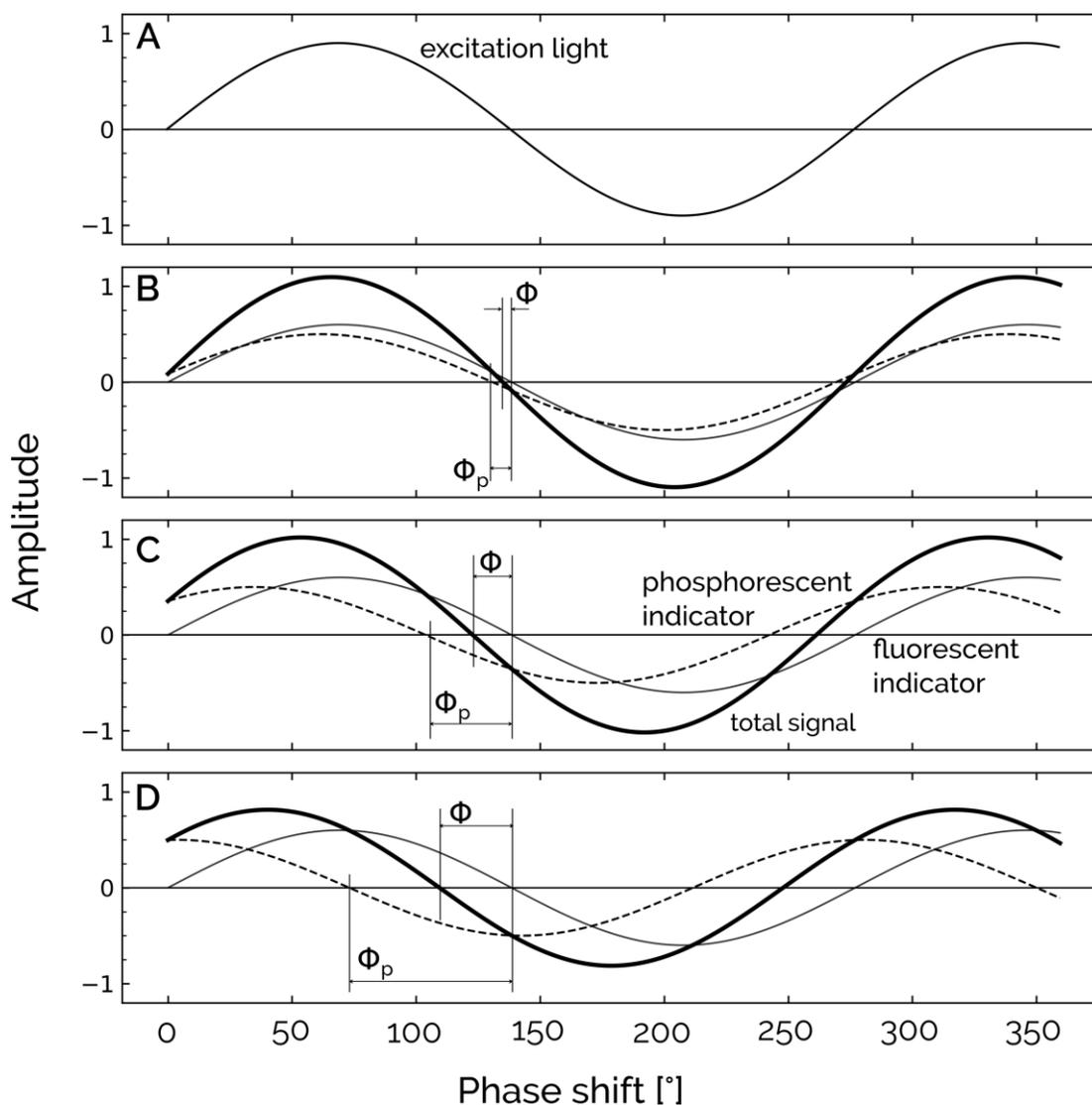
Figure 3-1: Schematics of the modified dual lifetime referencing method applied in the frequency domain redrawn from Borisov et al..[88] The phases of the short-lived fluorescent indicator Φ and the long-lived phosphorescent indicator Φ_p are marked as thin and dashed lines, respectively. The superimposed signal is shown as thick line. In subplot (**A**), the excitation light is shown as sinusoidally intensity-modulated light with a modulation frequency optimized to the decay time of the long-lived indicator dye. In the subplots below (**B-D**), the response functions of both individual indicator dyes and their superimposed signal at a certain intensity ratio are shown in three situations: In the subplot (**B**) the response functions are shown in the absence of any analyte, whereas in the subplots (**C**) and (**D**) the decay times of the fluorescent indicator or respectively of the phosphorescent indicator is reduced compared to (**B**).

However, for the modified dual lifetime referencing technique, the reference dye is substituted by another long-lived luminescent indicator sensitive to a second analyte. The long-lived indicator still acts as reference dye for the short-lived fluorescent indicator, however, the observed phase shift depends now on both analyte concentrations.[100] Therefore, two frequencies are required for signal deconvolution in order to precisely determine the concentrations of both analytes. A schematic of the modified dual lifetime referencing technique, enabling the simultaneous monitoring of two analytes, is shown in Figure 3-1. Upon simultaneous excitation of both indicator dyes by a sinusoidally intensity-modulated light at specific modulation frequency (see Figure 3-1A), both dyes

undergo demodulation and their response functions are delayed in time due to the respective analyte concentration as it is shown in Figure 3-1. The extent of the demodulation dm for each response function depends on the chosen modulation frequency and the decay time of the indicator dye. Its correlation is described by equation 3.4:

$$dm = \frac{1}{\sqrt{1 + (2 \cdot \pi \cdot f_{\text{mod}} \cdot \tau)^2}} \qquad \textbf{3.4}$$

with τ is the decay time of the indicator dye and f_{mod} is the modulation frequency. At low modulation frequencies in the range of few kHz and small decay times, the demodulation is almost 1 and the amplitude of the response function remains unchanged. However, instead of observing the individual phase shifts of each indicator, the superimposed (total) phase shift is determined combining the response functions of both indicators with respect to the specific amplitude ratio.

When only the fluorescent indicator responses to the analyte of interest, its emission intensity changes yielding an increased decay time while the decay time of the long-lived indicator dye remains unchanged as it is shown in Figure 3-1C. Following the principle of superposition, the overall phase shift and the total signal are slightly changed. The influence of the long-lived indicator dye is shown in Figure 3-1D. When the emission intensity of the long-lived indicator dye alters due to its response to the second analyte of interest, the total signal is more affected in terms of phase shift and demodulation. For signal deconvolution in order to extract information of both analytes of interest from the overall signal, two different modulation frequencies are required.

3.3.2. Mathematical Description of the modified Dual Lifetime Referencing

Since the excitation function is a sinusoidally intensity-modulated wave function, the response function is sinusoidally modulated as well and can be described in general terms by equation 3.5:

$$A_i(\varphi) = a_i \cdot \sin(\omega \cdot t + \varphi_i), \text{ with } \omega = 2 \cdot \pi \cdot f_{\text{mod}} \qquad \textbf{3.5}$$

where a_i is the amplitude of the individual indicator dye at the modulation frequency f_{mod} and φ_i is the specific phase shift of the indicator containing information about the analyte concentration.

Following the principle of superposition and assuming that the decay time of the fluorescent dye is too short for determination, the overall signal response function including their overall phase shift and their corresponding amplitude can be computed as follows:

$$A_{\text{total}} = \sqrt{a_1 + a_2 + 2 \cdot a_1 \cdot a_2 \cdot \cos(\varphi_2)} \qquad \textbf{3.6}$$

$$\tan(\Phi) = \frac{a_2 \cdot \sin(\varphi_2)}{a_1 + a_2 \cdot \cos(\varphi_2)} \qquad \textbf{3.7}$$

where a_1, a_2 are the amplitudes of the short-lived and the long-lived dye, φ_2 is the measurable phase shift of the long-lived dye and Φ is the overall phase shift of the dual optical sensor.

Transformation of equation 3.6 and substitution into equation 3.7 yield into the final description of the individual amplitude a_1 which only depends on the overall phase shift Φ, the amplitude of the long-lived dye a_2 and its corresponding phase shift φ_2:

$$a_1 = \left[\frac{\sin(\varphi_2)}{\tan(\Phi)} - \cos(\varphi_2) \right] \cdot a_2 \qquad \textbf{3.8}$$

While the amplitude undergoes demodulation and alters according to the chosen modulation frequency, the signal intensity remains constant. The amplitude of the individual indicator dye can therefore be converted into a constant signal intensity, when the actual demodulation factor is taken into account:

$$I_i = \frac{a_i}{dm_i(\tau_i, f_{\mathrm{mod}})} \qquad \textbf{3.9}$$

where the demodulation factor dm is determined according to equation 3.4 considering the actual decay time of the indicator dyes τ_i and their individual signal amplitudes a_i.

Since the signal intensity of the individual indicator dye remains constant for different modulation frequencies, the relative signal intensity of both indicator dyes is constant as well. A general expression of the intensity ratio is defined in equation 3.11 and can be simplified assuming that the demodulation factor of the short-lived indicator dye dm_1 is almost 1. In order to eliminate the incommensurable phase shift of the long-lived dye φ_2, a conversion of the trigonometric functions according to equation 3.10 is required, considering the correlation of the phase shift and the decay time defined by equation 3.2.

$$\frac{I_1}{I_2} = \frac{a_1 \cdot dm_2}{a_2 \cdot dm_1} \approx \frac{a_1 \cdot dm_2}{a_2} = \left[\frac{\sin(\varphi_2)}{\tan(\Phi)} - \cos(\varphi_2) \right] = \frac{2 \cdot \pi \cdot f_{mod} \cdot \tau_2 - \tan(\Phi)}{(1 + (2 \cdot \pi \cdot f_{mod} \cdot \tau_2)^2) \cdot \tan(\Phi)} \qquad \textbf{3.11}$$

$$\sin(\varphi_2) = \sin(\arctan(2 \cdot \pi \cdot f_{\mathrm{mod}} \cdot \tau_2)) = \frac{2 \cdot \pi \cdot f_{\mathrm{mod}} \cdot \tau_2}{\sqrt{1 + (2 \cdot \pi \cdot f_{\mathrm{mod}} \cdot \tau_2)^2}} \qquad \textbf{3.10}$$

While the signal amplitude of each indicator dye depends on the chosen modulation frequency, the signal intensity and therefore the intensity ratio remain constant. In the final form, the intensity ratio of the fluorescent dye I_1 compared to the phosphorescent dye I_2, only depends on measureable parameters, such as the overall phase shift of the dual optical sensor Φ at a specific modulation frequency f_{mod} and the decay time of the long-lived indicator dye. For signal deconvolution and calculation of the decay time of the long-lived indicator dye τ_2, the intensity ratio and corresponding operating parameters are determined at two different modulation frequencies f_1, f_2.

$$\tau_2 = \frac{f_1^2 - f_2^2 \pm \sqrt{(f_2^2 - f_1^2)^2 - 4 \cdot f_1 f_2 \cdot (f_1 \cdot \cot(\Phi_2) - f_2 \cdot \cot(\Phi_1)) \cdot (f_2 \cdot \cot(\Phi_2) - f_1 \cdot \cot(\Phi_1))}}{4\pi \cdot f_1 f_2 \cdot (f_1 \cdot \cot(\Phi_2) - f_2 \cdot \cot(\Phi_1))}$$

$$\textbf{3.12}$$

After calculation of the decay time of the long-lived indicator dye τ_2 and insertion into equation 3.11, the intensity ratio of the dual optical sensor can be determined. For the modified dual lifetime referencing technique, both intrinsic parameters are required in order to precisely determine the concentration of both analytes. Finally, it has to be highlighted that equation 3.12 is only valid for indicators undergoing a mono-exponential luminescence decay.[88]

3.3.3. Estimating the Relative Error

Errors will occur in quantitative analysis at least due to mistakes in observation and will increase due to the iterative character of various multivariate approaches, wherefore the estimation of the error propagation across the measurement and evaluation process is essential to consider. Thus, the reliability of the measurement results depends on the chosen algorithm and the individual results may differ.[82]

In general, an error is the difference between an individual measurement and its true value and is thus a measure of bias.[102] There are different types of errors ranging from gross errors to determinate and even indeterminate or random errors. While gross errors result from a carelessness and are thus avoidable, determinate errors are characterized as a systematic bias that affect the result of the individual measurement always in the same direction. Thus, determinate errors affect the accuracy of the final measurement result and may originate from measurement drifts, from an inhomogeneity or an impurity of the sample or may result from limitations of the measurement equipment. Besides the errors mentioned before, indeterminate errors are random errors affecting the precision of the measurement result. However, random errors can be reduced by an appropriate measure of spread.[102]

Since the true value of a sample is usually unknown, the anticipated value of the measurement and a reasonable measurement range of possible values, also known as measurement uncertainty, is provided. The measurement uncertainty is determined by the sum of all individual errors and its propagation across the evaluation procedure. Similar to the multiplicity of different types of errors, different approaches for the error analysis and the propagation of the measurement uncertainty exist. However, for estimating the maximum error of an individual measurement propagated across the evaluation procedure, the linear error propagation can be applied. Furthermore, if the error is small compared to the operating parameters, Taylor's series expansion can be applied for error analysis and the estimation of the measurement uncertainty.[103] For linear error propagation and estimation of the maximal error it is assumed that individual errors always sum up and a compensation of individual errors is therefore excluded (see equation 3.13).

$$f(x + \Delta x, y + \Delta y, \dots) = f(x, y, \dots) + \left|\frac{\partial f}{\partial x}\right| \Delta x + \left|\frac{\partial f}{\partial y}\right| \Delta y + \dots = f + |\Delta f| \qquad \textbf{3.13}$$

Applying the linear error propagation to the modified dual lifetime referencing technique, random errors can occur for the overall phase shifts and the amplitudes of the dual optical sensors arising from the background noise of the instrument and the reading inaccuracy. Thus, the equations for calculating the decay time of the long-lived indicator dye and the intensity ratio of the dual optical sensors must be updated, considering a linear error propagation. The error for the decay time $\Delta\tau$, calculated according to equation 3.12, arises from the measurement uncertainty in the phase shifts recorded at two different modulation frequencies Φ_1, Φ_2. The error in the decay time of the long-lived indicator dye is then used to calculate the corresponding error of the intensity ratio of the dual optical sensor following the definition of equation 3.11. Besides the propagated error of the decay time $\Delta\tau$, the measurement uncertainty in the recorded phase shifts affects the accuracy of the intensity ratio of the dual optical sensor. According to Taylor's series expansion, the estimation of the maximal error for the decay time of the long-lived indicator dye and the intensity ratio of the dual optical sensor is defined in equation 3.15 and equation 3.14:

$$\Delta\tau \;=\; \left|\frac{\partial\tau}{\partial\Phi_1}\right|\Delta\Phi_1 + \left|\frac{\partial\tau}{\partial\Phi_2}\right|\Delta\Phi_2 \qquad\qquad \textbf{3.15}$$

$$\Delta\frac{I_1}{I_2} \;=\; \left|\frac{\partial\frac{I_1}{I_2}}{\partial\tau}\right|\Delta\tau + \left|\frac{\partial\frac{I_1}{I_2}}{\partial\Phi}\right|\Delta\Phi \qquad\qquad \textbf{3.14}$$

In order to develop an optimized evaluation procedure for dual optical sensors based on the modified dual lifetime referencing technique, the estimation of the maximal error and a linear error propagation across the evaluation procedure was taken into account. For preliminary considerations, it was assumed that errors are of statistical nature arising only from the measurement uncertainty in the recorded phase shifts.

3.3.4. Implementation of Dual Optical Sensors

This subchapter is dedicated to the mathematical descriptions of the individual dual optical sensors realized within the dissertation based on the modified dual lifetime referencing technique. The considerations of the linear error propagation and the estimation of the maximal error were taken into account. The final simulation and evaluation procedures are combined and integrated within a graphical user interface (GUI), called dualSense, which was written in Python. A manual describing the operating instructions of dualSense is provided in part III - *Results targeting the Hybrid Sensing.*

3.3.4.1. Single Indicator Sensor

Indicator dyes that are sensitive to two different analytes at the same time, combining a prompt fluorescence as well as a delayed fluorescence in one single luminophore, are referred to as single indicator dyes. Those indicator dyes are preferred and provide a decisive advantage compared to multiple indicator dyes regarding their long-term stability, as these indicator dyes are less dependent on spectral alterations, such as bleaching or leaching. This can be explained by the dependence of the hybrid sensor on the specific intensity ratio of the short-lived and the long-lived moiety as it is described by equation 3.2. If both moieties are combined in the same luminophore, the effective intensity ratio remains unchanged, even if spectral alterations occur during ageing compromising the individual moieties to the same extent. Although those single indicator sensors are beneficial, they are generally rare and hard to synthesize. Within this dissertation however, we realized such a single indicator sensor for simultaneous determination of oxygen and temperature. The functional principle and the mathematical description are provided in the following subsection, while preliminary experiments of this sensor are described in the third part in chapter 9.

Functional principle

While it is possible to analyze alterations of different properties (e.g. refractive index, absorption, fluorescence or luminescence decay time), the functional principle of the single indicator sensor bases on alterations of spectral properties of the single indicator sensor in dependence on the oxygen content and the temperature of the sample. The sensitivity of the hybrid sensor bases on the dynamic quenching according to the two-site Stern-Volmer model as well as on the concept of thermally-activated delayed fluorescence (TADF).[104,105] As described in the previous chapter, the dual lifetime referencing technique analyzes the luminescence decay time as well as the intensity ratio of the hybrid sensor. The intensity ratio is analyzed in terms of prompt fluorescence compared to delayed fluorescence. While both analytes affect the luminescence decay time of the indicator as well as the intensity of the delayed fluorescence, the prompt fluorescence intensity is insensitive to both analytes and remains constant. When the oxygen concentration is increasing, the luminescence decay time decreases, following an exponential decay. At the same time, the delayed fluorescence intensity of the sensor decreases as well, while the prompt fluorescence intensity is independent of the oxygen content, resulting into an overall increase in the effective intensity ratio. On the other hand, when the temperature is increasing, the luminescence decay time of the sensor decreases, however in this case, following a linear correlation. Moreover, the intensity ratio of the prompt fluorescence intensity and the delayed fluorescence decreases with increasing temperature, as the delayed fluorescence intensity increases and the prompt fluorescence intensity remains constant. Thus, as both analytes affect both intrinsic spectral properties of the single indicator sensor in a distinct manner, we finally obtain a fully described system which can be solved following the mathematical approach described in the following subsection.

Mathematical description

Luminophores which combine both a prompt fluorescence and a delayed fluorescence in the same compound are straightforward in their implementation as dual optical sensor in terms of material input and production process. In contrast to their ease of production, the luminophores require more sophisticated instrumentations due to their luminescence decay time of few milliseconds. Thus, their modulation frequencies and the corresponding superimposed phase angles are also low and might suffer from electro-magnetic interferences. Therefore, the optical hybrid sensors and, in particular, their precision are highly prone to errors. Moreover, such single indicator hybrid sensors are hard to characterize, as the concentration of only one analyte should be varied, while the other must be kept constant over time. Thus, a preliminary characterization of the individual sensor before combining both sensors to a hybrid sensor is hardly possible. Consequently, the approach of choice was to generate a comprehensive calibration matrix for the intrinsic sensor parameters (luminescence decay time and intensity ratio), where both analytes are monitored at the same time.
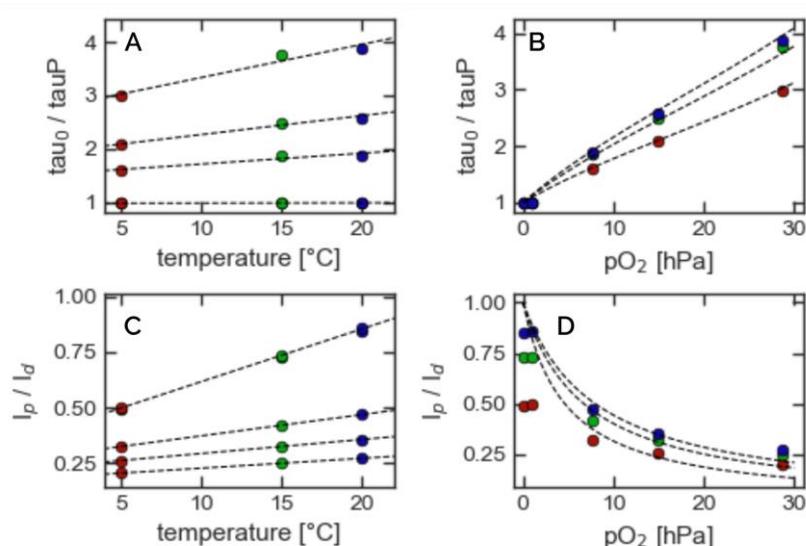
Figure 3-3: Relative luminescence decay time and intensity ratio of the single-layered hybrid sensor sensitive to oxygen and temperature. The intensity ratio is described in terms of the prompt fluorescence intensity I_p versus the delayed fluorescence intensity I_d. The temperature dependence of the decay time (**A**) and the intensity ratio (**C**) can be fitted by a linear regression, whereas the oxygen dependence (**B, D**) must be fitted by a two-site Stern-Volmer model.

Based on the calibration matrix, the sensor behavior was extrapolated to iteratively compute intermediate data points to further improve the resolution of the calibration matrix.

The single indicator sensor was implemented as dual optical sensor enabling the simultanuous determination of temperature and oxygen. A synoptic overview of the individual calibration measurements within an oxygen range of 0 to 30 hPa for three different temperatures is shown in Figure 3-3. Approximating the Arrhenius-model, the temperature dependence of the luminescence decay time follows a linear correlation within a temperature range of 5 to 30 °C.[88,106] In addition, it was found out that also the intenisty ratio of the hybrid sensor can be described best by a linear

correlation. Thus, the luminescence decay time or the relative decay time and the intensity ratio of the hybrid sensor can be expressed by equation 3.16.

$$y = k \cdot T + c \qquad\qquad \textbf{3.16}$$

where y is either the (relative) luminescence decay time of the hybrid sensor or its intensity ratio and k, c are the fitting parameters of the linear regression within a temperature range of 5 to 30 °C.

While a linear correlation can be approximated for the temperature dependence, a two-site Stern-Volmer model was used to describe the oxygen dependence of the relative luminescence decay time τ/τ_0 and the intensity ratio of the hybrid sensor $I_{\text{delayed}}/I_{\text{prompt}}$.[88,105] A general expression of the two-site Stern-Volmer model is shown in the following equation:

$$y = \frac{f}{1 + K_{sv,1} \cdot [O_2]} + \frac{1 - f}{1 + K_{sv,2} \cdot [O_2]} \qquad\qquad \textbf{3.17}$$

where y is either the relative luminescence decay time of the hybrid sensor or its intensity ratio at specific oxygen concentration $[O_2]$, f is a weighting factor describing the quenching distribution within the indicator dye and $K_{sv,1}, K_{sv,2}$ are Stern-Volmer quenching constants for the collisional and the static quenching. In its reduced form, the second Stern-Volmer constant can be expressed by the first invariable and a proportional factor $K_{sv,2} = K_{sv,1} \cdot m$.

For the dual optical sensor enabling the simultaneous determination of temperature and oxygen, both fitting equations were combined to a comprehensive calibration function in order to generate a comprehensive calibration matrix of the luminescence decay time and the intensity ratio. Within this calibration function, a temperature dependence was assumed for the Stern-Volmer constant and the weighting factor f.

$$y(T, [O_2]) = \frac{f(T)}{1 + K_{sv,1}(T) \cdot [O_2]} + \frac{1 - f(T)}{1 + K_{sv,2}(T) \cdot [O_2]} = \qquad\qquad \textbf{3.18}$$

$$= \frac{k_f \cdot T + c_f}{1 + (k_{Ksv} \cdot T + c_{Ksv}) \cdot [O_2]} + \frac{1 - k_f \cdot T - c_f}{1 + m \cdot (k_{Ksv\,2} \cdot T + c_{Ksv\,2}) \cdot [O_2]}$$

In equation 3.18 y is either the relative luminescence decay time or the intensity ratio of the hybrid sensor. To generate an appropriate calibration matrix of the dual optical sensor and to explicitly determine all six fitting parameters in equation 3.18, it might be reasonable to use at least six different calibration points with defined oxygen concentration and temperature.

For the analysis of an unknown measurement point based on the phase shifts at two different modulation frequencies, the oxygen concentration and the temperature are determined geometrically including their measurement uncertainty as it is shown in Figure 3-4. The evaluation bases on the assumption that the solution of a particular measurement point is clearly defined, since the junction of the intersection lines between two continuously and always monotonous planes is determined explicitly. The intrinsic parameters of the hybrid sensor are calculated with respect to

the equations 3.11 and 3.12. Subsequently, the junction of the calibration matrices, defined by equation 3.18, and the luminescence decay time or, respectively, the intensity ratio are computed yielding to an intersection line for both parameters. Subsequently, the point of intersection can be revealed, when both intersection curves are set equal to each other.
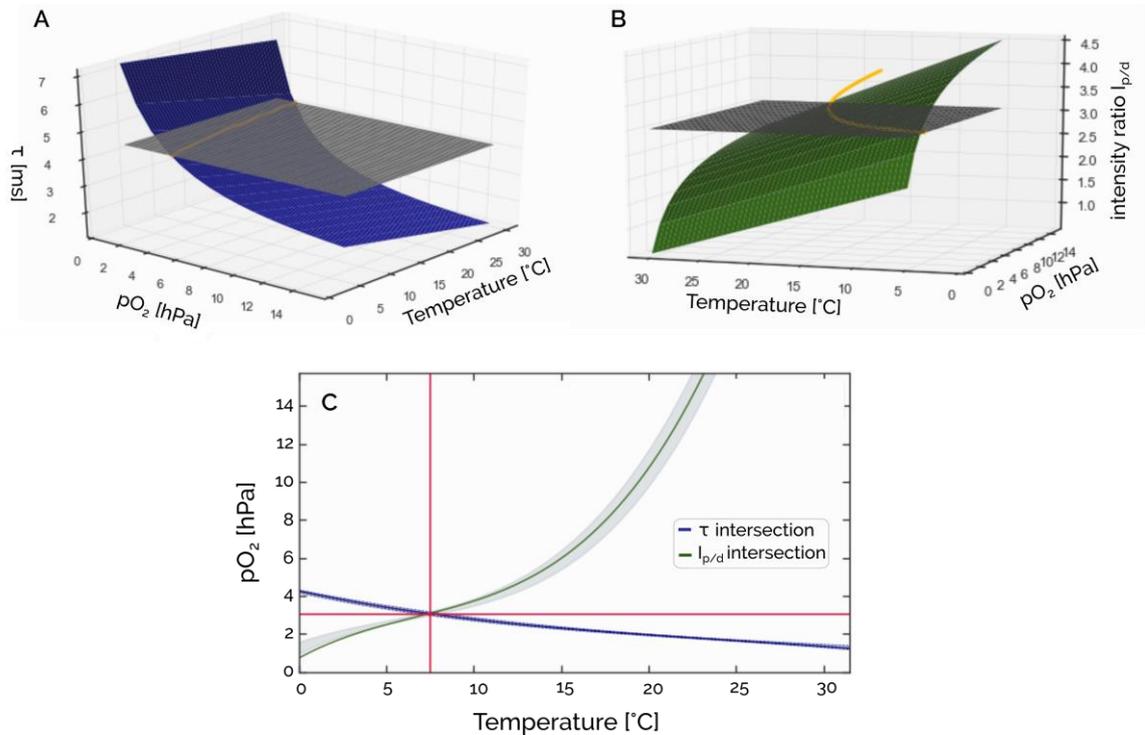
Figure 3-4: Analysis of an unknown measurement point for the single-layered dual optical sensor sensitive to oxygen and temperature. The oxygen concentration and the temperature at the measurement point are determined geometrically assuming that the junction of two intersection lines will reveal an explicitly defined point of intersection in order to determine the oxygen concentration and the temperature of the measurement point. First the intersection lines of the measurement point and the calibration matrix are computed for the luminescent decay time (**A**) and the intensity ratio (**B**) of the hybrid sensor. Both intersection lines are marked in yellow. Subsequently, the junction of both intersection lines is computed (**C**) to explicitly determine the oxygen concentration and the temperature of the measurement point. The point of intersection is marked as red reticule.

3.3.4.2. Two Indicator based Hybrid Sensors

While it is convenient to monitor multiple analytes by one single indicator dye, the characterization and calibration of single indicator based sensor systems is challenging. In particular, if different indicator dyes are incorporated within one sensor layer, the indicator dyes will interact and affect the overall sensor behavior in an indeterminate manner. Thus, it is more straightforward to combine multiple sensor layers that are sensitive to one specific analyte. The characterization of the individual sensor is thus less challenging and it is assumed that the individual sensors will not interfere, when the sensor layers are incorporated within the hybrid sensor.

For measurement evaluation according to the modified dual lifetime referencing technique presented before, the luminescence decay time of the phosphorescent indicator dye is calculated according to equation 3.12 using the overall phase shift at two definite modulation frequencies. The

first analyte content can subsequently be revealed with respect to the physicochemical principles of the respective analyte and with respect to an appropriate calibration. Based on the calibration conducted for the individual sensor dye, the phosphorescence intensity can be determined as well. The actual fluorescence intensity of the hybrid sensor can then be determined, when the actual intensity ratio is calculated according to equation 3.11. The fluorescence intensity is then again correlated according to the physicochemical principles of the second analyte of interest. Consequently, both analytes can successively be determined, when the physicochemical principles are applied to laboratory measurements and an appropriate calibration is conducted in advance. For prediction of the anticipated phase shift of the hybrid sensor at different modulation frequencies, the evaluation is reversed according to equation 3.2.

Hybrid Sensors for Simultaneous Determination of pH and Oxygen

Functional principle

As described in the previous chapter, the modified dual lifetime referencing technique substitutes the insensitive reference indicator by another long-lived indicator dye that is now sensitive to another analyte, but it still acts as reference for the short-lived indicator dye. Thus, the dual optical sensor combines two indicators within one sensor system enabling the simultaneous determination of two analytes. Using this concept, we addressed the disadvantage of many pH sensors which are cross-sensitive to dissolved oxygen and turned this challenge into a dual optical sensor that enables the determination of both analytes at the same time and at the same position. Again, the functional principle bases on changes in spectral properties in dependence on the oxygen content and the pH value. In this case, the long-lived indicator dye and, in particular, its luminescence decay time as well as its luminescence intensity are only sensitive to the oxygen concentration and base on a dynamic quenching according to the two-site Stern-Volmer model.[105] In contrast, the short-lived indicator is mainly sensitive to the pH of the sample and cross-sensitive to dissolved oxygen. The sensitivity of the pH indicator depends on the protonation and deprotonation of the indicator and is described by the Photoinduced Electron Transfer (PET) and the Intramolecular Charge Transfer (ICT) mechanism.[45] The description of these concepts is beyond the scope of the thesis, however, the key aspect of these mechanisms is that population of the energy level of the indicators and their relaxation processes are affected in a distinct manner by the named analytes. Thereby, the spectral properties, in particular the luminescence intensity as well as the luminescence decay time, change in dependence on the analyte concentration. While the correlation of the phosphorescence intensity is a function of the oxygen concentration and can be described by an exponential function, known as two-site Stern-Volmer fit, the fluorescence intensity follows a sigmoidal fit, which is normally referred to as Boltzmann fit. Moreover, it has to be taken into account that the short-lived indicator might also be affected by the actual oxygen content. As a consequence, for simultaneous determination of both analytes, the oxygen concentration must be calculated in the first place using the luminescence decay time or rather the actual phosphorescence intensity of the long-lived indicator of the hybrid sensor with respect to equation 3.12. After subsequent calculation of the

actual intensity ratio according to equation 3.11, the fluorescence intensity of the short-lived indicator can be determined. However, for precise determination of the actual pH value, the actual oxygen concentration and their influence on the pH calibration is taken into account by re-calibration of the corresponding Boltzmann parameters.

Mathematical description

The characterization of the hybrid sensor sensitive to pH and oxygen bases on the characterization of the individual sensor layers. The resulting fitting parameters are subsequently transferred to the hybrid sensor assuming that the individual layers do not interfere in other ways except it was previously defined.

While the oxygen dependence of the long-lived indicator dye is described best by the two-site Stern-Volmer equation according to equation 3.17, the pH dependence of the fluorescence intensity of the short-lived indicator dye and the overall phase shift follows a Boltzmann fit shown in equation 3.19.[107,108]

$$y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{\frac{pH - pK_a}{\text{slope}}}} \qquad \textbf{3.19}$$

where y can either be the fluorescence intensity I_F or the overall phase shift $\cot(\Phi)$; *Bottom* and *Top* are the minimum and maximum of the sigmoidal regression curve, pK_a is the point of inflection and *slope* refers to the actual slope of the Boltzmann fit with an anticipated value of 1. While the point of inflection and the slope of the Boltzmann fit are independent from the actual oxygen concentration, the other parameters are not. They must be extrapolated according to the preliminarily determined oxygen concentration.

For appropriate simulation of a measurement point, the oxygen sensor is iteratively calibrated, assuming that the fitting parameters are independent of the pH value. Further, the conversion of the decay time and the phosphorescence intensity is required and can be characterized in advance. The decay time of the long-lived indicator can be determined, as well as the corresponding phosphorescence intensity. The pH sensitive sensor is calibrated according to equation 3.19, where the minimal fluorescence intensity *Bottom* or respectively the minimal cotangent is substituted by the phase shift originated from the decay time of the oxygen sensitive indicator. Further fitting parameters are determined with respect to the actual oxygen concentration. Thus, the anticipated values for the overall phase shift of the hybrid sensor at different modulation frequencies as well as the corresponding intensity ratio can be computed.

For evaluation of an unknown measurement point, the procedure is reversed and the overall phase shifts are used to calculate the luminescence decay time of the oxygen-sensitive indicator dye and the intensity ratio of the hybrid sensor. Based on the actual decay time, the oxygen concentration can be calculated assuming a two-site Stern-Volmer model for fitting (see equation 3.17). In addition, the decay time is converted into the phosphorescence intensity considering a previously

determined conversion factor. Subsequently, the fluorescence intensity is determined with respect to the actual intensity ratio. Transformation of equation 3.19 allows then the determination of the actual pH value at a specific oxygen concentration.

The estimation of the maximal error originated from the measurement uncertainty in the phase shifts and its propagation across the iterative evaluation procedure is conducted with respect to section 3.3.3.

Hybrid Sensors for Simultaneous Determination of pH and Temperature

Functional principle

As almost all luminescent dyes display a temperature dependence, it might be elegant to design a hybrid sensor that enables an auto-correction of the actual temperature.[45] Thus, we conceived a two-indicator hybrid sensor enabling the determination of the pH value of the sample in consideration of the actual temperature. The functional principle of the hybrid sensor bases on changes in spectral properties in dependence on the actual temperature and on the pH value of the sample. For the temperature sensing, changes in optical properties, in particular in the luminescence intensity and the decay time, originate from the Boltzmann distribution and the temperature dependence of the electronic band structure of the indicator.[109] As mentioned previously, the measurement of the luminescence decay time – if possible – is preferred and can be described by the Arrhenius model.[110] For the pH sensing, changes in optical properties of the short-lived indicator mainly depend on the pH value of the sample and can be analyzed due to protonation and deprotonation of the indicator. Again, these effects are explained by the Photoinduced Electron Transfer (PET) and the Intramolecular Charge Transfer (ICT) mechanism which link changes in the luminescence intensity and the decay time to the actual pH value.[45] The response curve of the luminescence decay time, or rather the superimposed phase shift of the hybrid sensor, is described by the sigmoidal function of the Boltzmann fit. The fact that the overall phase shift is used explains the temperature cross-sensitivity of the Boltzmann parameters. Thus, for joint evaluation of both analytes, the temperature must be determined in the first place using the luminescence decay time of the long-lived indicator of the hybrid sensor with respect to equation 3.12. Then, the actual fitting parameters must be re-calibrated to the actual temperature enabling a temperature correction of the actual pH according to the Boltzmann equation.

Mathematical description

Combining the physicochemical principles of the previous hybrid sensors, it is straightforward to characterize the hybrid sensor and to implement an appropriate evaluation strategy. As before, to approximate the Arrhenius-model, a linear regression was assumed to describe the temperature dependence of the hybrid sensor between 0-40 °C. Here, a Boltzmann fit was used to describe the pH dependence of the overall phase shift of the hybrid sensor rather than its fluorescence intensity.[108] Unlike the previous hybrid sensor, all regression parameters exhibit a cross-sensitivity

to temperature, which is why the regression parameters of equation 3.19 must be compensated linearly and extrapolated to the actual temperature before the actual pH value can be determined.

After determination of the luminescence decay time according to equation 3.12, the actual temperature can be computed, when the slope and the abscissa of the linear regression curve defined in equation 3.16 are calibrated with respect to two calibration points. Then, the actual pH value can be calculated using the overall phase shift and its pH correlation. To reveal the actual pH value, the equations 3.19 and 3.2 are equated with each other as it is shown below and all regression parameters of the Boltzmann sigmoid are temperature compensated:

$$\cot(\Phi) = \cot(\varphi_P) + \frac{1}{\sin(\varphi_P)} \cdot \frac{A_F}{A_P} \cong \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{\frac{pH - pK_a}{\text{slope}}}} \qquad \textbf{3.20}$$

$$pH = pK_a(T) + slope(T) \cdot \log\left(\frac{I_P \cdot dm_P}{I_F} \cdot [\text{Top}(T) \cdot \sin(\varphi_P) - \cos(\varphi_P)] - 1\right) \qquad \textbf{3.21}$$

where Φ is the superimposed phase shift of the hybrid sensor and φ_P is the phase shift of the long-lived temperature-sensitive indicator, which can be calculated from the phosphorescence decay time according to equation 3.3. A_F/A_P is the ratio of amplitudes of the hybrid sensor that can be calculated from the intensity ratio, when the demodulation factor is taken into account. Since the minimum of the Boltzmann sigmoid *Bottom* corresponds to the phase shift of the phosphorescence indicator, the equation can be simplified. Substitution and transformation allows thus the direct calculation of the pH value according to equation 3.21 after temperature compensation of the relevant regression parameters:

Similar to the previous dual sensors, the estimation of the maximal error originated from the measurement uncertainty in the phase shifts is taken into account with respect to section 3.3.3.

Hybrid Sensors for Simultaneous Determination of Carbon Dioxide and Oxygen (C)O2

Functional principle

The functional principle of the hybrid sensor sensitive to carbon dioxide and oxygen is the same as it was for the hybrid sensor sensitive to oxygen and pH. The principle bases on specific changes in spectral properties in dependence on the oxygen content and the carbon dioxide value of the sample. Again, the long-lived indicator dye and, in particular, its luminescence decay time as well as its luminescence intensity rely on the concept of dynamic quenching following the two-site Stern-Volmer model.[105] However, the short-lived indicator is sensitive to carbon dioxide and, to some extent, cross-sensitive to the oxygen content of the sample. Similar to a pH indicator, the sensitivity of the carbon dioxide indicator depends on protonation and deprotonation of the indicator and is described by the Photoinduced Electron Transfer (PET) and the Intramolecular Charge Transfer (ICT) mechanism.[45] Those concepts are used to link changes in spectral properties to the analyte

concentration. On the one hand, the phosphorescence intensity is a function of the oxygen concentration described by the exponential function of the two-site Stern-Volmer model.[105] On the other hand, the fluorescence intensity is primarily affected by the carbon dioxide concentration and follows an exponential function as well. However, it has to be taken into account that the short-lived indicator might be cross-sensitive to the actual oxygen content of the sample. Therefore, for simultaneous determination of both analytes, the oxygen concentration must be calculated first, using the luminescence decay time of the long-lived indicator of the hybrid sensor with respect to equation 3.12. Then, the actual intensity ratio of the hybrid sensor is determined according to equation 3.11. The actual fluorescence intensity of the short-lived indicator can then be determined. However, for precise determination of the effective carbon dioxide concentration, the actual oxygen concentration and its influence on the calibration of the short-lived indicator is taken into account by re-calibration of the corresponding Stern-Volmer parameters.

Mathematical description

Similar to the previous hybrid sensors, the dual optical sensor targeting the simultaneous monitoring of carbon dioxide and oxygen was characterized individually assuming that the sensor layers will not interfere in other ways than previously defined, when they are incorporated in the dual optical sensor.

While the regression of the oxygen-sensitive indicator dye was straightforward assuming a two-site Stern-Volmer correlation following equation 3.17, no correlation was implemented for the fluorescent indicator dye sensitive to carbon dioxide at the first place. However, upon analysis of various regression approaches, it was investigated that the two-site model works best to describe the carbon dioxide dependence of the fluorescence intensity exhibiting an optimal goodness of fit in terms of a minimal chi-square value (Figure 3-5).

Consequently, after calculation of the luminescence decay time of the oxygen-sensitive indicator dye, the relative decay time was converted into the corresponding phosphorescence intensity. The actual oxygen concentration is determined after a two-point calibration (ambient air and an anoxic calibration point). In addition, the intensity ratio is used to determine the actual fluorescence intensity of the hybrid sensor, which is then again used to iteratively calculate the actual carbon dioxide concentration according to equation 3.17 with respect to two calibration points. In order to take into account that the overall phase shift of the hybrid sensor is affected by the carbon dioxide concentration, the Stern-Volmer regression parameters for carbon dioxide calculation are approximated via re-calibration at minimal phosphorescence intensity. However, it has to be highlighted that this approach assumes no cross-sensitivity of the oxygen sensor to the actual carbon dioxide concentration.
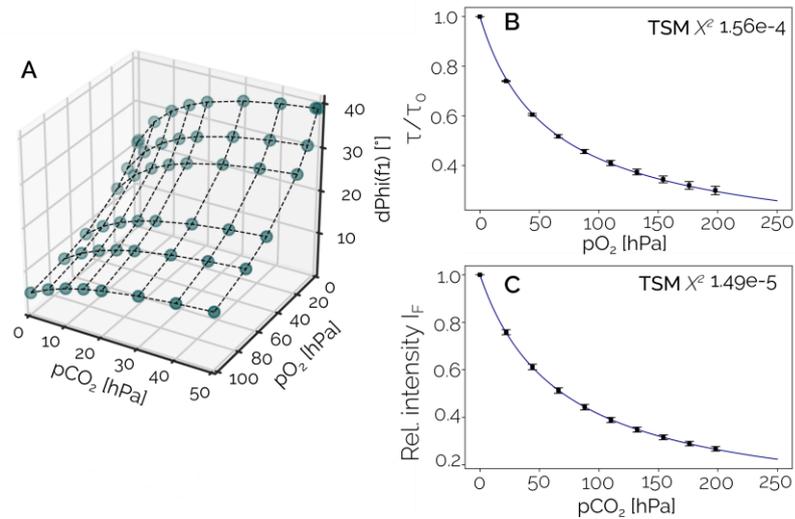
Figure 3-5: Schematics for the CO2/O2 hybrid sensor based on the modified DLR principles. (**A**) The superimposed phase shift was simulated for a modulation frequency of 3 kHz, while the calibration plots (**B-C**) represent measurement data. For both individual sensors, their dependence is well described by the two-site Stern-Volmer model exhibiting an optimal goodness of fit in term of a minimal χ2-value of 1.56e-4 (pO2, **B**) and 1.49e-5 (pCO2, **C**), respectively.

Part II

Results targeting the Algae Monitoring

4. Compact and Low-Cost Fluorescence Based Flow-Through Analyzer for Early-Stage Classification of Potentially Toxic Algae and in Situ Semi-quantification
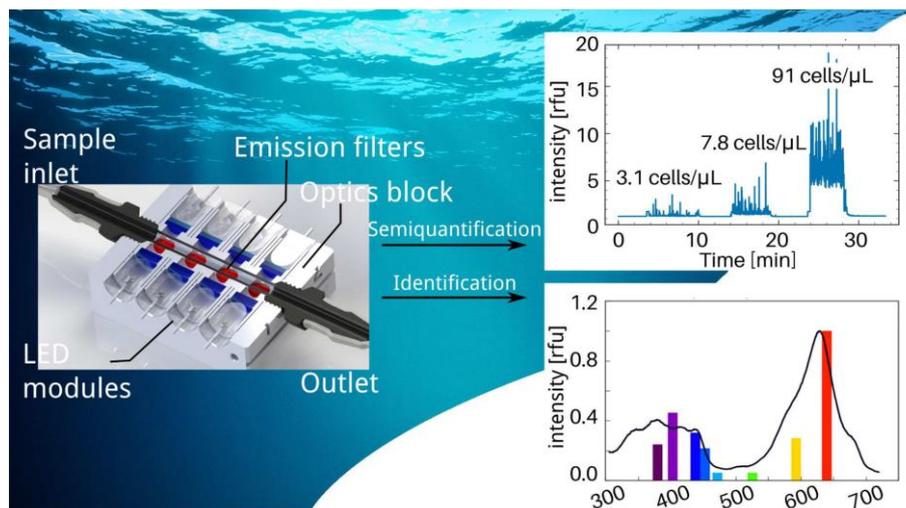
4.1. Preface for the Manuscript

This publication represents the development and evaluation of the compact multi-wavelength algae detection module, named ALPACA. The ALPACA features a miniaturized, modular and cost effective design resulting in a highly versatile and flexible monitoring tool. Integrated into submersible housings, the ALPACA enables the continuous classification of marine phytoplankton in coastal environments and the early-stage identification of algal species potentially able to produce biotoxins with a detection limit of 10 cells per liter.

The ALPACA was developed as part of the EU-funded project SCHeMA aiming at the integration of multiple sensing systems in one single submersible housing for in situ high resolution mapping of various anthropogenic and natural chemical compounds[5,111]. The design of the algae detection module was mainly specified by the submersible housing which was developed in cooperation with Idronaut S.r.l.. In the following course of the dissertation, the prototype of the ALPACA was intended to use as basic building block for the multiparametric sensing system.

This publication mainly focusses on the fabrication, the electrical and optical components used in the detection device as well as on the verification of the systems performance. Furthermore, internal calibration and standardization strategies were developed during the thesis to standardize and correct the fluorescence signal recorded on the ALPACA aiming at greater inter-comparability within the detection system and between experiments.

Additionally, we implemented a data evaluation strategy enabling the autonomous classification of algae in mixed samples as well as the control of biomass alterations. For the classification algorithm, we investigated the characteristic spectral properties of eight main algal classes and identified eight LEDs that are considered to be significant for reliable differentiation of the algal samples. For further details on the theoretical background, please refer to chapter 2. The results of the study about spectral properties and the evaluation strategy were published elsewhere (see chapter 5) as well as they were incorporated into a software interface for data evaluation (see chapter 6).

Compact and Low-Cost Fluorescence Based Flow-Through Analyzer for Early-Stage Classification of Potentially Toxic Algae and in Situ Semi-quantification

Authors: Silvia E. Zieger[1*], Günter Mistlberger[1], Lukas Troi[1], Alexander Lang[1], Fabio Confalonieri[2] and Ingo Klimant[1]

[1] Institute of Analytical Chemistry and Food Chemistry, Graz University of Technology, Stremayrgasse 9, 8010 Graz, Austria

[2] Idronaut S.r.l., Via Monte Amiata 10, 20861 Brugherio (MB), Italy

*Corresponding author: zieger@tugraz.at

4.2. Abstract

The occurrence and intensity of (harmful) algal blooms (HABs) have increased through the years due to rapidly changing environmental conditions. At the same time, the demand for low-cost instrumentation has increased substantially, enabling the real-time monitoring and early-stage detection of HABs. To meet this challenge, we have developed a compact multi-wavelength fluorometer for less than 400 USD. This is possible by using readily available and low-cost optical and electronic components. Its modular design results in a highly versatile and flexible monitoring tool. The algae detection module enables a continuous identification and control of relevant algal groups based on their spectral characteristics with a detection limit of 10 cells per liter. Besides its usage as a benchtop module in the laboratory, the algae module has been integrated into submersible housings and applied in coastal environments. During its first in situ application in the port of Genoa, seawater samples of mixed algal composition were used to demonstrate the successful discrimination of cyanobacteria and dinophytes as well-known toxin producing classes. Fabrication, operation, and performance as well as its first in situ application are addressed.

4.3. Introduction

Harmful algal blooms (HABs) encompass phytoplankton species of undefined concentration, which may cause harm to the surrounding environment, public health or to economic and ecological structures and functions.[10] Harmful blooms of dinophytes and cyanobacteria are known to produce biotoxins that accumulate through the food web, leading to various health risks for animals and humans, such as shellfish poisoning.[4,16,112–114] Besides, blooms can further cause harm to fishes, marine mammals or co-occurring organisms due to the biomass, they achieve and hypoxia from their decay.[11–13] Although there exists no general threshold level above which an algal bloom is recognized as harmful, the WHO proposes a guidance value during recreational exposure to cyanobacteria with 20×10^3 cells per milliliter.[115] Due to their potential threats even at low concentration, the awareness of HABs has been extended and the demand to understand bloom dynamics and related biological processes has increased.[3,114,116,117] Consequently, monitoring tools are essential for real-time characterization of rapidly changing conditions and episodic alterations in marine life. Continuous observation and early-stage identification of bloom compositions are crucial and preferred instead of time, energy and cost-intensive laboratory work which may miss episodic algal events.[4,10,118]

Detecting and tracking algal dynamics have been of great interest for marine biologists, and different approaches for characterization exist, for example remote sensing systems via satellites, high performance liquid chromatography (HPLC) analyses, microscopic evaluations or fluorometric analyses.[10] In particular, fluorometric measurements are sensitive methodologies that have been valuable in advancing the understanding of the distribution and composition of phytoplankton assemblages. In situ fluorometers are effective instruments to quantify phytoplankton in situ.[119–122] Their application has been increased even more since their accuracy and reliability have been improved and low-cost devices have been commercialized.[123,124] However, marine fluorometers have drawbacks concerning their restricted discrimination capability of mixed algal assemblages.[15,125] Moreover, marine fluorometers face problems with low algal concentrations and their sensitivity to ambient light.[123] On the other hand, (micro)flow cytometers are powerful tools for quantifying and characterizing taxonomic composition of individual cells.[10] However, (micro)flow cytometers suffer from high manufacturing cost and significant power requirement. In addition, their usage on small boats is limited due to their large size.[10,126] Moreover, high sensitivity and characterization opportunities come along only with significantly slower and more complicated systems.[122] A combination of both, a miniaturized, low-cost fluorometer, not only feasible for quantifying but also for classifying algal assemblages, will be a valuable contribution to marine science.

The general approach for classifying algae, a multivariate data evaluation based on chemo-taxonomic principles, has proven valuable in flow cytometry.[127,128] Phytoplankton groups vary in their characteristic pigment compositions and their spectral properties are defined by the presence / absence and relative composition of their accessory pigments. Therefore, a comprehensive measurement of the excitation properties of the phytoplankton enables their

indirect differentiation and identification at the order level after complex, multivariate discrimination analysis.[125,129,130]

It is crucial to identify and discriminate relevant algal phyla among other dominant groups in an algal assemblage at an early stage and across broad temporal and spatial scales. To face these challenges, we developed a compact and low-cost multi-wavelength fluorometer, which focusses on the excitation characteristics of the phytoplankton. The concept is summarized in its name – the Advanced Luminescence-based Phytoplankton Analysis and Classification Appliance (ALPACA). It is a sensitive and robust monitoring device enabling the real-time discrimination of relevant phytoplankton groups combined with an approximation of the cell density. Its response linearity is comparable to other commercial and non-commercial fluorometers for algae detection. However, its main benefit is the early stage classification of algal groups in mixed assemblages targeting the identification of well-known toxin producing algae (cyanobacteria and dinoflagellates). In this work, we describe design, fabrication and operation of the ALPACA. Moreover, we demonstrate its applicability in laboratory and in field measurements.

4.4. Materials and Methods

4.4.1. Chemicals and Materials

The optics block of the ALPACA, made out of a polyoxymethylene copolymer (POM-C), was designed in house and fabricated by protolabs (Germany, www.protolabs.de). A circular quartz capillary, 1.94 mm inner diameter and 40 mm long, was purchased from Hilgenberg (Germany, www.hilgenberg-gmbh.de). The microcontroller Olimex PIC32-PINGUINO-MICRO was obtained from Olimex Ltd (Bulgaria, www.olimex.com) and the circuit board was purchased from Ätzwerk (Germany, www.aetzwerk.de). The amplifier MAZeT MTI04 is from ams Sensors Germany GmbH (Germany, www.mazet.de) and the analogue-digital converter (ADC ADS1115) from Texas Instruments. The light-emitting diode (LED) driver (TLC5917) and transducer (RS485) were purchased from Farnell (Austria, www.at.farnell.com). All multi-layer capacitors, resistors, connectors and wires were also ordered from Farnell. The peristaltic pump, Minipuls 3, and its PVC tubing are from Gilson International (USA, www.gilson.com). The 375 nm UV-LED and the colored LEDs (405 nm, 430 nm, 450 nm, 475 nm, 525 nm, 590 nm and 640 nm) were obtained from Roithner Lasertechnik (Austria, www.roithner-laser.com). Metal housings for the excitation source, made from aluminum, were designed and fabricated in house. Four silicon PIN photodiodes (BPW34) were obtained from Vishay (Farnell, www.uk.farnell.com). Two types of longpass filters (RG-665; blocking range $\lambda_s = 650$ nm and RG-9; blocking range $\lambda_s = 710$ nm) and four bandpass filters for the excitation sources (BG-25, BG-39, F-39 BrightLine HC and F-49 ET) were obtained from bk Interferenzoptik (Germany, www.interferenzoptik.de) and AHF Analysentechnik (Germany, www.ahf.de), respectively. The plastic filter 'Primary Red' was obtained from LEE Filters

(UK, www.leefilters.com). Rhodamine 101 and Sulforhodamine 101 were purchased from SigmaAldrich (Austria, www.sigmaaldrich.com). Rhodamine 101 was diluted in ethylene glycol to a final concentration of 12 mM, whereas sulforhodamine 101 was diluted in different concentrations between 0.03–19.7 μM using water as solvent.

4.4.2. Algal Samples and Cultivation

A total of 18 different algal cultures were either purchased from the Culture Collection of Algae (University of Göttingen, Germany) or obtained by the French Research Institute for Exploitation of the Sea (IFREMER, France). Each culture was maintained in conical flasks with f/2 growth medium at a salinity of 33 with silicate as appropriate and additional trace metals according to Guillard and Ryther.[131,132] The cultivating temperature was 19 °C and the irradiance with a cool white fluorescent tube was 37 μmol/(s m[2]). The light:dark cycle during cultivation was 10:14 h. Further information about the cultures is supplied in Supporting Information Table S4-1.

4.4.3. Instrument Development

4.4.3.1. Design and Fabrication

The design of the ALPACA has been developed in the frame of the FP7-SCHeMA project (2015).[5] ALPACA is constructed as a miniaturized multi-wavelength fluorometer operating in a modular way to enable an easy adjustment for different requirements and applications. The main part consists of an optics block machined from a polyoxymethylene copolymer with a geometrical dimension of 28 x 35 x 16 mm[3]. Inside, a circular quartz capillary is inserted, where the sample is pumped through during the measurement. Along this capillary, eight excitation channels – four at each side – and four emission channels are aligned at right angles. A schematic of the ALPACA is shown in Figure 4-1 and Figure S4-1 of the Supporting Information. According to the geometrical dimensions and the device configuration, the calculated volume of one measurement channel is 6 μL. Besides miniaturization, ALPACA aims at increasing the sensitivity and selectivity compared to common in situ fluorometers. Therefore, decrease of background effects and reduction of potential interferences between the measurement channels were crucial targets within the development process.
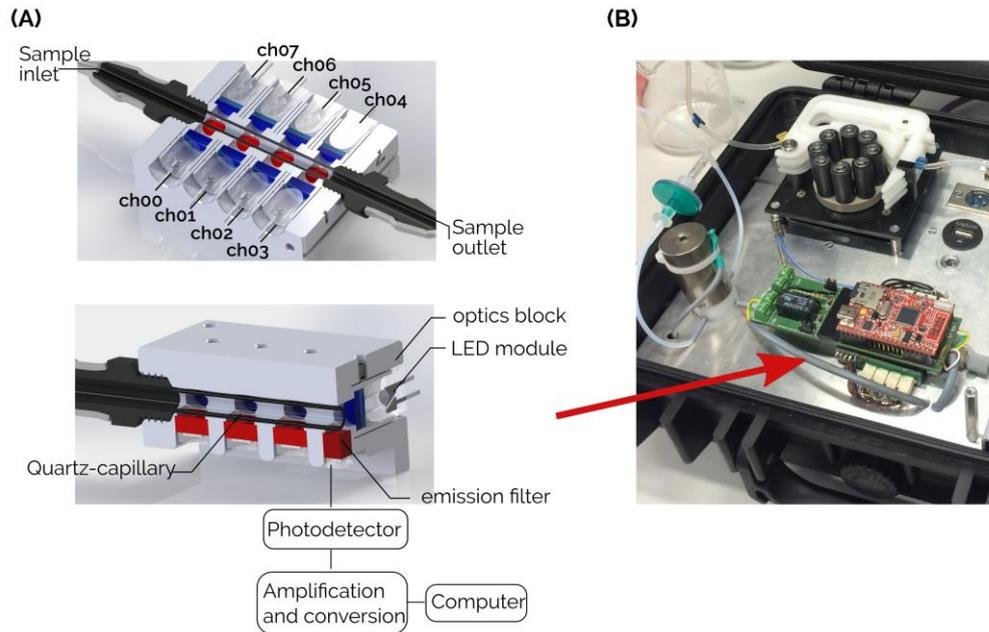
Figure 4-1: Schematic of the optical, electronics and fluidics setup of the ALPACA in top view and cross section. The outer dimensions of the ALPACA are 28.0 x 34.8 x 16.8 mm^3 (width x length x height). (**A**) In the cross section below, the arrangement of electrical components for signal detection, amplification and conversion is included. (**B**) Realized prototype of the miniaturized and multi-wavelength ALPACA.

Table 4-1 Spectral characteristics for all measurement channels of the ALPACA[a]

	Excitation wavelength (nm)	Detection range (nm)	Correction factors exc. @ 50 mA	Correction factors emission
ch00	438	650–800	1.00	4.6
ch01	453	650–800	0.62	4.6
ch02	472	650–800	0.49	4.6
ch03	640	710–800	0.49	1.0
ch04	403	710–800	1.72	1.0
ch05	380	650–800	0.92	4.6
ch06	593	650–800	0.08	4.6
ch07	526	650–800	0.38	4.6

[a] Besides the excitation wavelengths and the detection ranges, the correction factor (excitation and emission) for each measurement channel is also given. All photodiodes were covered with the longpass emission filter RG-665 expect the photodiode for channel 3 and channel 4. The recorded output signal is given in nanowatts.

4.4.3.2. Electrical and Optical Components

For the excitation of the phytoplankton within the measurement channels, different LEDs in the wavelength range between 380–640 nm are used. The LED selection aims to excite characteristic pigments within the light-harvesting complexes of the phytoplankton, allowing the identification of relevant algal phyla. No LED above 650 nm was selected for financial reasons and as it does not improve the identification efficiency of the ALPACA. The LEDs are combined with optical glass and

interference filters, in order to define the emission characteristics of the excitation source distinctively and to narrow their emission spectra further. Both, LED and filters, are placed in a metal housing for an easy replacement (Supporting Information Figure S4-2). Due to the metal enclosure design, the excitation source can be adapted to fit any 5 mm LED and excitation filters of any thickness between 0.2 and 4.0 mm. Further information about electro-optical characteristics of LEDs and filter combinations is provided in Table S4-2 of the Supporting Information.

The chlorophyll fluorescence emitted from the phytoplankton upon excitation, is recorded by four silicon PIN photodiodes, sensitive to visible light and near-infrared radiation. These photodiodes are covered by two different types of longpass filters, either an RG-665 or an RG-9 filter, in order to enable a detailed examination of the fluorescence originating from photosystem-II and photosystem-I, respectively. Using the longpass filter RG-665, fluorescence emitted from the biomass is recorded above 650 nm, whereas the other filter enables the fluorescence detection above 710 nm. These emission filters are additionally covered by a plastic filter to avoid auto-fluorescence of the emission filters and therefore enhance the signal-to-background ratio (SNR). All LEDs and emission filters are chosen to avoid interferences between excitation and emission channels. Moreover, possible interferences, e.g. from colored dissolved organic matter, yellow substances such as humic matter or suspended particles are attenuated by the mounted filters.

The optics block is connected to the fluidic system using opaque, pressure-stable fluidic connectors, in order to ensure a light-proof and pressure-stable system. Moreover, to prevent undesirable electromagnetic interferences, the whole system is mounted in a grounded metal enclosure.

For data acquisition and device control, a printed circuit board (PCB) was manufactured. Surface-mounted-device components (SMD components) are used to enable a miniaturized design of the electronics fitting the small optics block. The readout of the recorded fluorescence signals is done by a 15-bit Analog-Digital-Converter with a multichannel transimpedance amplifier (MAZeT-TIA). An Arduino-like board with an 80 MHz and 32-bit microcontroller is used for instrument control. The software interface is written using Python, an open source programming language.[133] The microcontroller enables a fast sequential measurement of the LEDs with a sampling rate of 264 Hz per channel. Additional light- / dark measurements for noise reduction and calibration of electrical components are implemented, reducing the effective sampling rate to 88 Hz per LED. Furthermore, the intensity of each LED, the amplification and the measurement frequency of the photodiodes can be regulated. The LED intensity can be linearly adjusted between 30–50 mA and is normally used at maximum current. The raw data can either be stored on an SD card or, especially during laboratory applications, transmitted via USB-port directly to a computer for external data evaluation.

The power consumption of the ALPACA in measurement mode, using an LED intensity of 50 mA, is 80 mA @ 5V. Therefore, the USB connection to the external computer can serve as power supply for the detection unit.

Besides the stand-alone application, the compact design of the ALPACA (width/length/height: 80/60/30 mm³ including electronics) makes it suitable for simple integration in submersible housings. Furthermore, an RS-485 communication port is implemented for communication and data transmission to an external host system.

4.4.3.3. Device Calibration

For correction of excitation sources, the relative LED intensity and all measurement channels are calibrated separately against an internal quantum counter (Rhodamine 101 in ethylene glycol, 12 mM). The concentrated dye solution exhibits a fluorescence quantum yield of 100% independent from the excitation wavelength.[44] Upon excitation, the actual fluorescence intensity of rhodamine 101 recorded on the ALPACA is compared to the theoretic signal intensity according to equation 4.1 (Figure S4-3 of the Supporting Information). For correction of the photodiodes, the fluorescence intensity of sulforhodamine 101 (19.7 μM) was compared against one another and for internal calibration. The correction factors are multiplied during evaluation with the recorded light intensity, resulting a corrected fluorescence emission signal of the phytoplankton. The correction factors, listed in Table 4-1 and Table S4-3 of the Supporting Information, are included in the data evaluation step to ensure inter- and intra-comparability of the measurement at any time. Recalibration is only necessary if the device setup is modified on the excitation or emission side.

$$\kappa(\text{rfu}) = \frac{I_{\text{theoretic}}\ (\text{nW})}{I_{\text{actual}}\ (\text{nW})}$$

4.1

4.4.3.4. Device Operation, Data Acquisition and Processing

The sample solution is pumped through the detection unit using a peristaltic pump at a maximum flow rate of 1.5 mL/min. Pump velocity and data acquisition rate ensure that each cell event is recognized at each measurement channel with at least three measurement points. This strategy is crucial to enable and improve individual signal analyses of mixed algal samples. In general, the operating current of the LEDs is set to 50 mA and the transimpedance of the amplifier is set to 20 MΩ, however, both can be regulated according to the cell density of the sample. Photoemission above 650 nm or 710 nm – depending on the measurement channel – is recorded by the photodiodes and converted into light intensity in nanowatt. The recorded and corrected light intensity over time is displayed in real-time. It equals the fluorescence intensity emitted from the phytoplankton upon excitation at eight distinct wavelengths (Figure 4-2A). Further information about the conversion is given in the Supporting Information.

For evaluation, blank and baseline corrections are processed using either seawater (laboratory use) or 70% ethanol (during field tests to avoid biofouling). Cell events are evaluated and the average light intensity, which is equivalent to the average fluorescence intensity at each excitation channel, is calculated and displayed as a histogram (Figure 4-2B). These steps are automatically processed

by the software; the user only has to define the evaluation range (Figure 4-2A). After characterization of more than 53 different phytoplankton using the ALPACA, we found that the emission ratio between the excitation at 438 nm (chlorophyll-*a*) and 526 nm (phycobiliproteins) is most suitable for a rough classification. For the empirical determination of threshold levels, we used phytoplankton of nine different classes under different light conditions and in different concentrations: Cyanobacteria, Rhodophyta, Dinophyta (Dinoflagellates), Bacillariophyta (Diatoms), Haptophyte, Chlorophyte, Ochrophyta, Cryptophyta and Euglenophyta. Here, a ratio below 0.69 indicates a dominance of cyanobacteria or rhodophytes, whereas a ratio above 0.83 indicates a dominance of other algae phyla. If the ratio of the two signals lies in between those thresholds, cyanobacteria and / or algae, the assignment is equivocal and both groups might be present in the sample. The result of this immediate / preliminary evaluation is displayed directly on the screen. For an identification of the algal class in detail, the normalized fluorescence intensity is evaluated using a pattern recognition algorithm. The multivariate discriminant analysis, here the linear discriminant analysis (LDA), compares the normalized and standardized fluorescence intensity of the sample relative to a set of reference patterns of known phytoplankton groups.[134] The aim is to examine the probability of group membership for the sample returning the algal class, where the measured fluorescence pattern fits best. An example of the data evaluation steps is shown in Figure 4-2. Further information about the statistical principle and the software is given in the Supporting Information. Furthermore, a pursuing article about the spectral characterization and the mathematical algorithm is currently in preparation.

4.4.4. Performance Evaluation

4.4.4.1. Algal Characterization and Comparison to Reference Fluorometer

For investigation of the normalized fluorescence pattern and spectral properties of the algae, fluorescence excitation- and emission-spectra (FEEMs) were acquired for each algal sample with a fluorescence spectrofluorometer (Fluorolog-3, Horiba Jobin-Yvon, France) at room temperature. The samples were held in suspension with glycerin to avoid sedimentation during measurement. 3D spectra were recorded from 300–750 nm excitation and 400–950 nm emission in order to reveal the fluorescence of chlorophyll-*a* and auxiliary pigments of the light-harvesting complex. Excitation spectra were compared to the normalized fluorescence pattern gained from the ALPACA to evaluate the reliability of the measurement results.

(A) Emission signal over time

(B) Fluorescence pattern / excitation spectra

(C) Probabilty of group membership

(D) 2D-Score plot
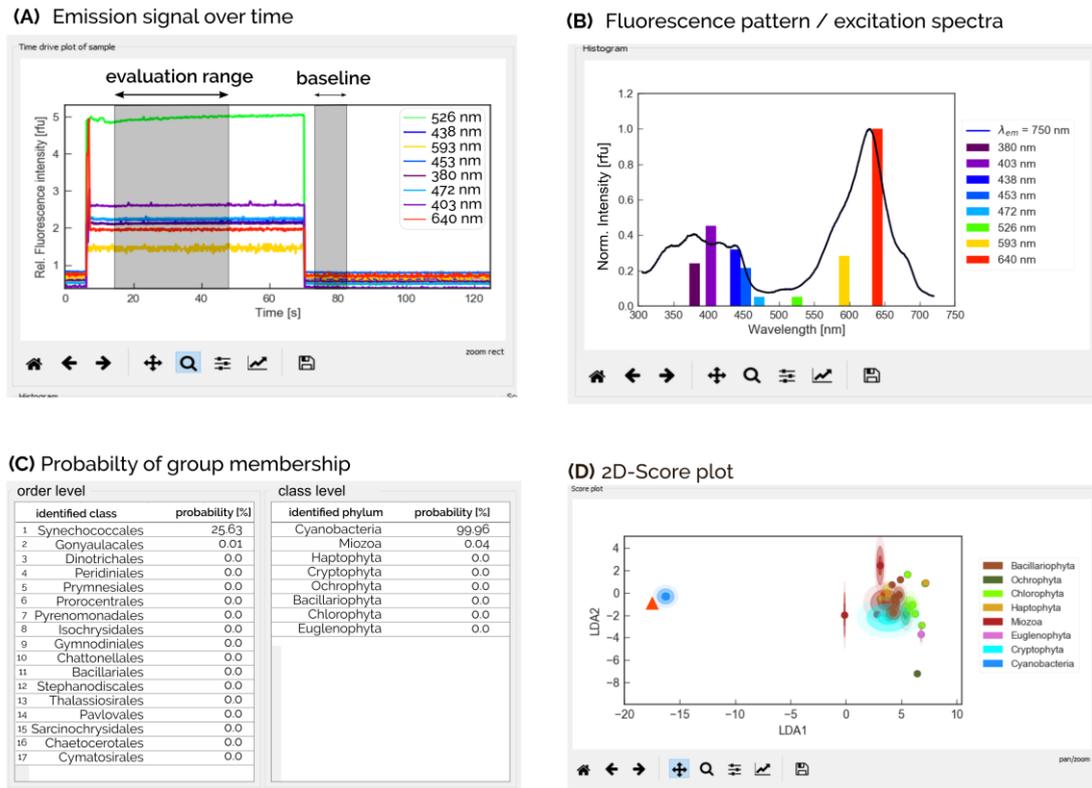
Figure 4-2: Fluorescence intensity upon excitation at eight different excitation wavelengths over time of a highly-concentrated cyanobacteria suspension (Synechococcus sp.) pumped through the detection module (**A**). In the software interface, the user defines the time range for evaluation and baseline correction (marked in grey). After processing the average fluorescence pattern at different excitation wavelengths (**B**), the linear discriminant analysis returns the probability of class membership for the unknown sample (**C**) and the score plot of the sample (**D**).

4.4.4.2. Linearity of the ALPACA and Limit of Detection (LOD)

The linearity of each measurement channel as well as the detection limit were investigated using $0.03-7.4\,\mu$M sulforhodamine 101 as fluorescence standard solution. The fluorophore was enclosed into capillaries for reasons of reproducibility and easy replacement. A blue 472 nm-LED was used for excitation and all data were blank-corrected. Furthermore, the cell density of three different algae species (*Amphora* sp*., Hemiselmis cf rufescens* and *Cyanobacteria* sp.) was correlated with the fluorescence intensity emitted upon excitation at 453 nm and 526 nm. Before measuring, algal samples were adapted to darkness to avoid any influence of light.[54,135] For determination of the cell density, an inverted light microscope equipped with eyepiece 10X and objective 20X (200X magnification, Carl Zeiss) was used according to Andersen.[136]

The limit of detection (LOD) is defined as the lowest cell density that can be detected by the ALPACA. However, the chlorophyll content of the algae depends, among other things, on the biovolume and consequently on the species.[137] Therefore, it was necessary to investigate the minimal cell density for different algae and cyanobacteria varying in cell size between $14-200$ μm².[138,139] The minimal cell density was then combined with the average measuring volume to calculate the LOD of the ALPACA.

4.5. Results and Discussion

4.5.1. Sample Analysis and Comparison to a Reference Fluorometer

Spectral differences between phytoplankton species belonging to different algal groups were investigated with the spectrofluorometer as shown in Figure S4-4 (Supporting Information). Based on these findings, different LEDs were selected as excitation sources for the ALPACA, as listed in Table S4-2 of the Supporting Information. Photosystems of cyanobacteria, which are dominated by phycobiliproteins, can mainly be excited in the near-infrared wavelength range. While phyocerythrin absorbs mainly between 545–565 nm and 498 nm, phycocyanin absorbs at ~ 620 nm.[69] Consequently, a green or orange LED in this specific wavelength range, enables the excitation of phycobilins and can be used for the reliable differentiation of cyanobacteria and rhodophyta from other algae. However, photosystems of algae are dominated by chlorophyll, which is present not only in the antenna complex, but also in the reaction center. Chlorophyll-*a* and its derivatives absorb maximally at their Soret peak at ~ 430 nm but also at higher wavelengths at 680 nm (Qy Band).[41] Therefore, a blue LED can be used to excite chlorophyll in order to discriminate phytoplankton from most other marine particles. Additional LEDs in the blue wavelength range are used to excite the most prominent, diagnostic pigments for a distinct discrimination of further phytoplankton classes.

In addition, the correlation of the normalized fluorescence patterns recorded either with the spectrofluorometer or with the ALPACA is successfully demonstrated (Figure S4-4c–d, Supporting Information). Apart from LED 593 nm, which in general exhibits quite low signal intensities, the discrete fluorescence pattern gained from the ALPACA coincides well with the reference spectra.

4.5.2. Linearity of the ALPACA and Limit of Detection (LOD)

The linearity of the ALPACA is successfully demonstrated using sulforhodamine 101 as fluorescence standard solution between 0.03–7.4 μM (spectral properties shown in Supporting Information Figure S4-7). The fluorescence intensity upon excitation with a blue 472 nm LED was blank-corrected. The coefficient of determination for each measurement channel is calculated to 99.35–100%. (Supporting Information Figure S4-8).

Semi-quantification of the biomass with monitoring tools in real-world applications is crucial to manage algal blooms. To demonstrate the semi-quantitative capability of the ALPACA, the cell density of different dark-adapted algal species was correlated against the fluorescence intensity emitted upon excitation at 453 nm (chlorophyll-*a*) or 526 nm (phycobiliproteins) (Supporting Information Figure S4-9 to Figure S4-11). The calculated coefficients of determination (R^2) indicate a linear correlation (Table 4-2). These coefficients of determination (R^2) are competitive with other commercial and non-commercial fluorometers for algae detection reported by the Alliance for

Coastal Technologies (ACT).[140] Furthermore, Figure 4-3a describes how the fluorescence signal depends on the biomass within the measurement chamber. At lower biomass, signal spikes of individual cells enable a detection of single cells and therefore a cell count for normal-sized algae and filamentous cyanobacteria. Experiments prove that single cells can be detected on the ALPACA with a signal-to-background ratio of ~ 3 (Supporting Information). Since the average measurement volume for one experiment is 100 mL, the detection limit (LOD) can therefore be calculated to 10 cells per liter. However, the detection limit increases for smaller cells, particularly for cryptophytes and unicellular cyanobacteria, when a single cell detection is not possible (Table 4-2).
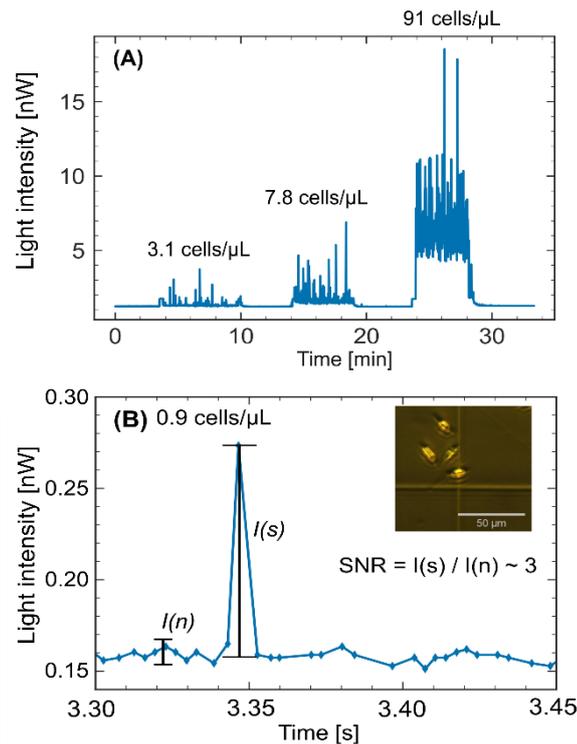
Figure 4-3: Variations of the corrected fluorescence signal of the diatom *Amphora* sp. depending on the biomass between 3.1–91 cells per microliter upon excitation at 453 nm. The photoemission is recorded as sum signal above 650 nm. (**A**). When the biomass in the measurement chamber is high, the average signal level is heightened, but gets resolved at lower cell densities (3.1 cells per microliter). When the cell density is even lower, i.e. 0.9 cells per microliter, individual signal spikes correspond to individual cell events passing through the measurement channel, allowing a single cell detection and cell counting (**B**). The slight increase of background signal due to multiple scattering effects within the capillary cannot be completely prevented.

Table 4-2 Limit of detection and coefficient of determination calculated after correlation of the cell density and emitted fluorescence intensity at 453 nm or 526 nm.

algal species	R² at 453 nm (%)	R² at 526 nm (%)	LOD (cells per liter)
Amphora sp.	99.78	99.32	10
Hemiselmis cf rufescens	99.47	99.44	20
Cyanobacteria sp.	97.89	98.59	10

4.5.3. Statistical Classification and Quality Criteria

Statistical quality criteria are used in order to ensure comparability amongst different evaluation strategies. Besides sensitivity and specificity, as typical statistical measures of performance, miss rate and accuracy are determined. These statistical measures were determined for two evaluation strategies: (1) for the preliminary classification method, using two LEDs for classification (438 nm / 526 nm), which are summarized in Table 4-3 and (2) for the discriminant analysis (LDA), which are summarized in Table 4-3. First the group membership was predicted using these strategies, then its accuracy was validated.

Sensitivity (equation 4.2) or true-positive rate (TPR), describes the proportion of positive and correctly identified events (T_P) compared to the amount of events that are effectively positive ($T_P + F_N$).

$$\text{sensitivity} = \text{TPR} = \frac{T_P}{T_P + F_N}$$

4.2

When using LDA, the sensitivity for cyanobacteria is 50%, due to a large spectral overlap of cyanobacteria and rhodophyta, however increases to 100%, if both classes are combined ('Cyanobacteria plus Rhodophyta'). However, using the preliminary evaluation with two LEDs, cyanobacteria and rhodophytes are combined as one group, obtaining an overall sensitivity of 100%. The sensitivity for reliable identification of dinophytes is 100%, using LDA.

The miss rate or false-negative rate (FNR) in equation 4.3 describes events which are positive in fact, but not identified as such (F_N) in relation to the amount of events that are positive in fact ($T_P + F_N$).

$$\text{FNR} = \frac{F_N}{T_P + F_N}$$

4.3

The miss rate is also known as the 'power' of an algorithm, whereby a low rate is preferred. Miss rate and sensitivity add up to 100%. For the preliminary evaluation, using two LEDs (438 nm / 526 nm), the miss rate is optimal with 0% for both groups, for 'Cyanobacteria plus Rhodophyta' and for the group 'other algae'. Using the discriminant analysis, the miss rate is also excellent, with 0% for 'Cyanobacteria plus Rhodophyta' and for dinophytes. The miss rate for the cyanobacteria alone, however, is reduced to 50%.

Specificity, also called true-negative rate (TNR), is the number of events that are correctly identified as negative in comparison to the number of events that are negative in fact ($T_N + F_P$).

$$\text{specificity} = \text{TNR} = \frac{T_N}{T_N + F_P}$$

4.4

The specificity is also called true-negative rate and it is excellent for the preliminary analysis. As shown in Table 4-3, the specificity for the whole group of 'Cyanobacteria plus Rhodophyta', but also for the cyanobacteria alone, is 100%. However, the specificity for the dinophytes is reduced to

86.7%, as more species, especially haptophytes, are classified as dinophytes (F_P). A reason for this cross-interference is the spectral overlap of haptophytes and dinophytes.

Accuracy, also called confidence level, describes events that are correctly classified compared to all predictions made (parent population) by the algorithm.

$$ACC = \frac{T_P + T_N}{T_P + F_P + T_N + F_N}$$

4.5

In this study, the accuracy for the preliminary analysis is optimal with 100%. Also for the LDA, the accuracy is optimal with 94.1% and 88.2% for cyanobacteria and dinophytes, respectively, while the common accuracy for 'Cyanobacteria plus Rhodophyta' is 100%. Due to the spectral overlap of dinophytes and haptophytes, the false-positive classification is increased, which therefore leads to a reduced accuracy for dinophytes. The false-negative classification is 0%. The sensitivity and accuracy for cyanobacteria emerge from the high spectral overlap between cyanobacteria and rhodophytes. This interference has to be reassessed and the analysis of more species might improve the sensitivity and discrimination efficiency.

In summary, accuracy and specificity are adequate and the false-negative rate is reasonable. A detailed overview of all discriminated algal groups for both evaluation steps and their quality criteria is shown in the Supporting Information Table S4-4.

Table 4-3 Quality criteria for the evaluation of 18 algal samples belonging to eight different algal phyla using two different evaluation strategies: (1) a rough analysis with 2 LEDs (438 nm / 526 nm) for the preliminary discrimination between 'Cyanobacteria plus Rhodophyta' and other phytoplankton phyla (preliminary evaluation), or (2) using LDA for the identification of further algal groups, in particular dinophyta.[a]

evaluation	algal group		sensitivity (TPR) (%)	miss rate (FNR) (%)	specificity (TNR) (%)	accuracy (ACC) (%)
preliminary	Cyanobacteria	plus	100.0	0.0	100.0	100.0
	Rhodophyta					
	other algae		100.0	0.0	100.0	100.0
LDA	Cyanobacteria		50.0	50.0	100.0	94.1
	Cyanobacteria	plus	100.0	0.0	100.0	100.0
	Rhodophyta					
	Dinophyta		100.0	0.0	86.7	88.2
	Other algae		50.0	50.0	100.0	82.4

[a] The quality criteria are calculated according to equations 4.2–4.6.

4.5.4. Field Samples and Validation

The long-term stability of the ALPACA and its analytical performance were evaluated during a field campaign at the port of Genoa (CNR station: 44°23'46.6"N, 8°55'51.6"E, Supporting Information Figure S4-13) between February 14 – March 13 2017. For this purpose, the ALPACA, which was calibrated in advance, was incorporated into a pressure-stable submersible housing.[5] A solution (70% ethanol) was included in the submersible probe enabling a blank measurement before each

measurement circle. The submersible probe was installed at a depth of 4.6 meters under the surface throughout the entire period, including small spatial variabilities due to tides. Seawater samples were collected for 10 minutes continuously at two-hourly intervals resulting in a 15 mL sample volume. This volume was evaluated on site by the ALPACA. In order to avoid potential risk of biofouling during the field campaign, the blank solution was retained in the fluidic system, whenever no measurement was executed. The average fluorescence intensity at each excitation channel was recorded to analyze alterations in algal composition and relative algal content over time. Measurement data were transmitted via a network controller to the marine station. On February 21 and February 28 2017, two water samples, each being 1 L, were collected at the surface at the CNR station and preserved in formalin for off-line validation. An inverted light microscope equipped with Oculars 10X and Objective 40X (400X magnification, Carl Zeiss) was used to determine the average algal composition and cell density of these samples following the Utermöhl method described by Hasle and Zingone et al.[141,142]

The fluorescence pattern recorded by the ALPACA is shown in Figure 4-4A. The fluorescence intensity decreased slightly at the end of the field campaign. Further, based on the fluorescence intensity, variations in the average algal composition and content were investigate. Findings were validated by microscopic analysis, which determines that the cell density was maximal with 21.4×10^3 cells per liter on February 21 and decreased more than half to 8.7×10^3 cells per liter one week later (February 28). Consequently, these results confirm the findings shown in the intensity plot over time (Figure 4-4A). In addition, the average algal composition for each sampling day was analyzed by the ALPACA applying the discriminant analysis. During these days, the ALPACA determines the relative algal content for dinophytes with 1.2−7.1% including natural fluctuations. These results are confirmed by the microscopic analysis which calculates the relative algal content for dinophytes of 4.2% (Figure 4-4B−C).
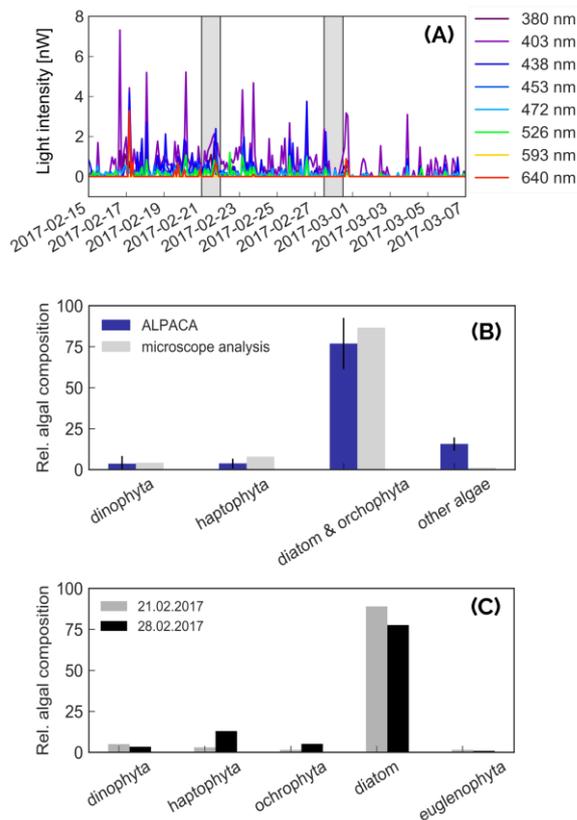
Figure 4-4: Results of the field campaign between February 14 – March 13 2017 with a continuous sampling frequency of two hours. (**A**): Average fluorescence pattern over time of the deployed ALPACA. Surface water samples of 15 mL were taken at two-hourly intervals and evaluated autonomously on site. Timeslots, when samples were taken for reference analysis, are marked in grey. (**B**): The average algal composition was analyzed on the ALPACA using measurements only from the two validation dates (21 and 28 February). The results were obtained by applying the linear discriminant analysis. Furthermore, the results were compared to the off-line results determined manually by a taxonomic expert under the microscope. In case of the ALPACA, the unspecified group 'other algae' also contains cyanobacteria which were not identified by the expert. (**C**): Deviation of the algal composition and content counted under the microscope for validation purpose. The total biomass decreases from 21.4×10^3 cells per liter on 21 February to 8.7×10^3 cells per liter on 28 February.

4.5.5. Strengths and Limitations of the ALPACA

During the evaluation procedure in coastal environment, ALPACA has shown a stable and robust performance with a technology readiness level (TRL) 7. The prototype performance was successfully demonstrated in an operational environment, although it is far from commercialization. Its modular design and small footprint make it attractive for various applications beyond laboratory use, for example on boats, moorings or for incorporation into submersible housings. Furthermore, the device is built up from only low-cost components, which reduces the material costs below 400 USD. Finally, the algae detection module operates at a low energy consumption (80 mA @ 5 V for measurements with 50 mA LED intensity). After measurement and autonomous data processing procedure, the device is switched off completely. The combination of these factors with a very simple device operation makes the ALPACA useful for small agencies, universities for educational purposes and scientific surveys. Due to the measurement in flow-through mode and reduced dependence on ambient light, the sensitivity of ALPACA is enhanced compared to commercially available fluorometers. Another strength of the detection module is the early stage identification of at least 1 cell / measurement volume (6 μL) for algal cells. The resulting detection limit of 10 cells per milliliter is below the proposed guidance value of the WHO (20×10^3 cells per milliliter), allowing an early-stage detection algal blooms before they might cause harm to their surroundings. In addition, it was successfully demonstrated that the biomass correlates to the recorded light

intensity with similar performances than other fluorometers for algae detection reported by the ACT. Therefore, a rough estimation of the phytoplankton density and the control of biomass can be provided by the ALPACA. However, its main benefit compared to other fluorometers, is its capability to describe bloom compositions in real-time by (multivariate) analysis of eight excitation wavelengths. It was shown with in situ analysis, that the autonomous evaluation of fluorescence patterns and, in particular, a distinct identification of dinophytes and cyanobacteria is reliable with a confidence level of 94.1% (cyanobacteria) or 88.2% (dinophytes), respectively. In addition, other algal groups can be discriminated according to the complexity of the underlying training database.

Despite all the advantages mentioned above, the ALPACA is currently only suitable for coastal regions in surface water as the pigmentation of the phytoplankton vary under different environmental conditions. The main reason for this limitation is the missing training databases and tests with species from other habitats. This limitation does not arise from technical restrictions, but from operational reasons as the characteristic pigment pattern of the phytoplankton species depends on its habitat. However, surface water comprises a large range of environments in coastal regions. After appropriate investigation of known phytoplankton species and, if required, slight adaption of the system, ALPACA can be applied in further regions. Moreover, the multivariate discriminant analysis relies on an appropriate training databases of known phytoplankton species and their spectral properties. In order to enhance the discrimination capability of cyanobacteria and rhodophytes or haptophytes and dinophytes, respectively, it might be suitable to strengthen the training database further. Nevertheless, the first evaluation of statistical measures in this study indicates a good performance of the ALPACA.

Acknowledgements

4.6. Supporting Information

4.6.1. Algal Characterization

Cell shape and size were investigated under the inverted light microscope. The algal samples were fixed with Lugol solution and placed in a counting chamber, which was used as reference to determine the cell length. Pictures were taken with a SensiCam from pco (200X magnification) and ImageJ was used to determine the average cell size of the phytoplankton samples.

Table S4-1 Taxonomic and structural characteristics of all phytoplankton species used to assess the performance of the ALPACA.* Species unknown; taxonomic description only at order possible

Phylum	Class	Order	Species	Shape and cell size
Cyanobacteria	*Cyanophyceae*	*Cyanobacteria*	unknown*	long, filamentous chains without nucleus; 4.2 x 111.1 μm
		Synechococcales	*Synechococcus* sp.	rod-shaped and nonflagellated; 6.8 x 8.3 μm
Rhodophyta	*Florideophyceae*	Bonnemaisoniales	*Bonnemaisonia hamifera*	medium-sized feathery or tufted form; multicellular
	Rhodellophyceae	Glaucosphaerales	*Rhodella maculata*	spherical; surrounded by a plasma membrane; radius 11.6 μm
Dinophyta (Dinoflagellates)	*Dinophyceae*	*Gonyaulacales*	*Alexandrium minutum*	spherical to ellipsoidal, armoured; 27.6 x 26.8 μm
		Peridiniales	*Heterocapsa triquetra*	two joined cones with unilateral protrusion; 20.1 x 16.5 μm
		Prorocentrales	*Prorocentrum cassubicum*	ovoid shaped with small perforation at the center and thickened margins; 18.9 x 31.3 μm
Cryptophyta	*Cryptophyceae*	*Cryptomonadales*	*Hemiselmis cf rufescens*	cone with half sphere, flagellated; 8.8 x 7.9 μm
Haptophyta	*Coccolithophyceae*	*Coccolithales*	*Pleurochrysis elongata*	spherical to drop-shaped; 17.1 x 17.4 μm
		Isochrysidales	*Isochrysis galbana*	ellipsoid, bi-flagellate; 9.6 x 8.5 μm
Ochrophyta	*Eustigmatophyceae*	*Eustigmatales*	*Eustigmatos magnus*	spherical, uniflagellate, eyespot at one end; radius 10.5 μm
Bacillariophyta (Diaotm)	*Mediophyceae*	*Thalassiosirales*	*Skeletonema costatum*	cylindrcal cells forming marginal ring of spines; 7.6 x 9.2 μm
	Bacillariophyceae	*Naviculales*	*Phaeodactylum tricornutum*	fusiform; 4.8 x 11.1 μm
	Bacillariophyceae	*Thalassiophysales*	*Amphora* sp.	convex dorsal and straight ventral margin; clear raphe with radial dorsal striae; 7.3 x 20.8 μm
Chlorophyta	*Chlorophyceae*	*Chlamydomonadales*	*Chlamydomonas reinhardtii*	oval, bi-flagellate; 9.9 x 13.2 μm

			spherical to spindle-shaped, eyespot on one side, unicellular, bi-flagellate; 9.4 x 13.3 μm
		Dunaliella salina	spherical to spindle-shaped, eyespot on one side, unicellular, bi-flagellate; 9.4 x 13.3 μm
	Sphaeropleales	*Scenedesmus* sp.	rotational ellipsoid; 7.3 x 13.7 μm
Trebouxiophyceae	*Chlorellales*	*Chlorella vulgaris*	spherical; 7.4 x 11.1 μm

4.6.2. Signal Conversion into Light Intensity

Photodiodes record the fluorescence signal emitted in Volts from the phytoplankton samples upon excitation. Taking the internal resistor of the amplifier (R = 20 MΩ) and the spectral sensitivity of the photodiodes (0.5 W/A at 700 nm) into account, we can convert the signal current into light intensity according to

$$1\text{V} \cdot \frac{1}{20\ \text{MΩ} \cdot 0.5\text{A/W}} =\ 100\text{nW} \qquad\qquad \textbf{S 4.1}$$

4.6.3. Finalized Prototype of the ALPACA Included in the Measurement Housing

Figure S4-1: Final laboratory setup for the ALPACA. During measurement, the algae module (**1**) is enclosed in a metal housing (**2**) to prevent interferences from ambient light or electromagnetic field (EMF) radiation. The valve on the inlet (**3**) changes between an optional filter inlet for rinsing and the sample inlet. The pump (**4**) is placed behind the outlet to ensure continuous flow.

4.6.4. LED and Filter Combinations
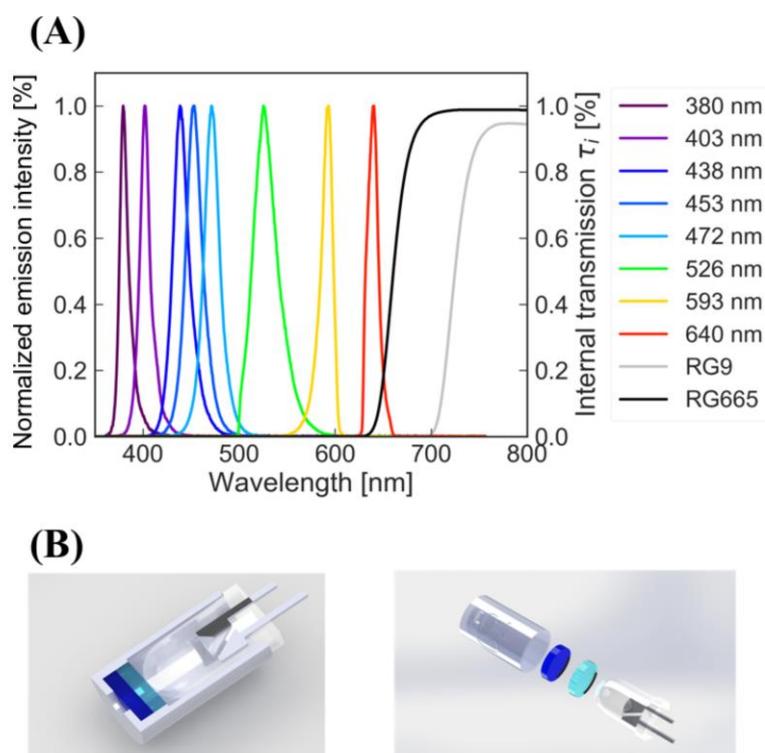
Figure S4-2: Spectral properties of the LEDs combined with optical glass and interference filters for distinct excitation of photosynthetic pigments and the long pass emission filters (RG665 and RG9) used in the ALPACA (**A**). LEDs and filters are embedded in a metal housing for easy replacement (**B**).

Table S4-2 Electro-optical characteristics of LEDs used in the ALPACA and LED- filters combination used in the LED module for pigment excitation.

Peak wavelength λ_P (nm)	Half width $\Delta\lambda$ (nm)	Viewing angle $2\Theta_{1/2}$ (deg)	Forward current I_F (mA)	Radiated power I_R (mW)	Optical filter
380	10	± 8	20	2.5	BG-25 / BG-39 2 + 2 mm
403	15	± 8	20	1.5	BG-25 / BG-39 2 + 2 mm
438	20	± 10	20	22	BG-25 / BG-39 2 + 2 mm
453	25	± 8	20	20	BG-25 / BG-39 2 + 2 mm
472	20	± 15	50	370	BG-25 / BG-39 2 + 2 mm
526	30	± 8	20	42	F-39-558 / BG-39 2 + 2 mm
593	30	± 15	70	641	F-39-558 / BG-39 2 + 2 mm
640	20	± 8	20	51	F-49-645 1.12 mm

4.6.5. Device Calibration
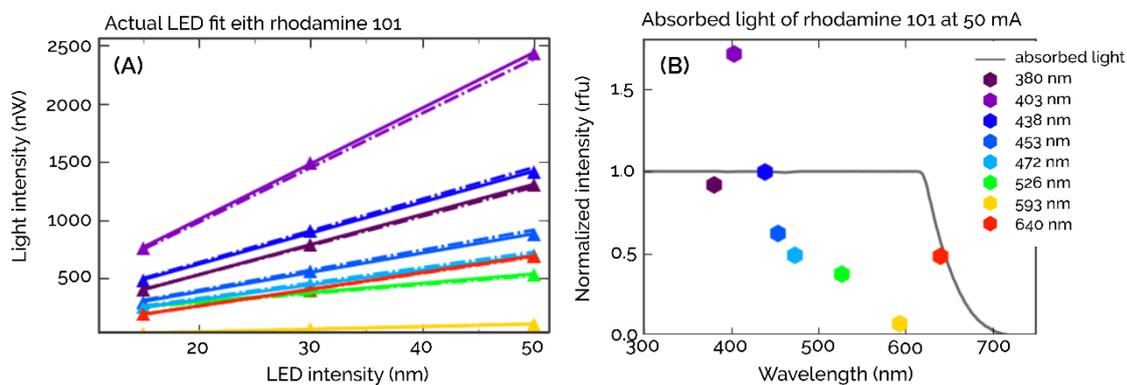
Figure S4-3: Principle of LED calibration: (**A**) To determine the correction factors k for each measurement channel, the actual fluorescence intensity of the internal quantum counter (rhodamine 101) was recorded upon excitation with an LED at three different LED intensities to calibrate the ALPACA along all possible amplifications. (**B**) The actual fluorescence intensity is then compared to the theoretical signal intensity.

Table S4-3 Spectral properties of the LEDs and their correction factors k measured within the detection module against an internal quantum counter (rhodamine 101). The correction factors are used for the calibration of the excitation sources.

Peak wavelength λ_P (nm)	Half width $\Delta\lambda$ (nm)	Correction factors k against rhodamine 101 @ 50 mA (1)
380	10	0.921
403	15	1.718
438	20	1.000
453	25	0.622
472	20	0.492
526	30	0.379
593	30	0.075
640	20	0.488

4.6.6. Sample Analysis and Comparison to Reference a Fluorometer
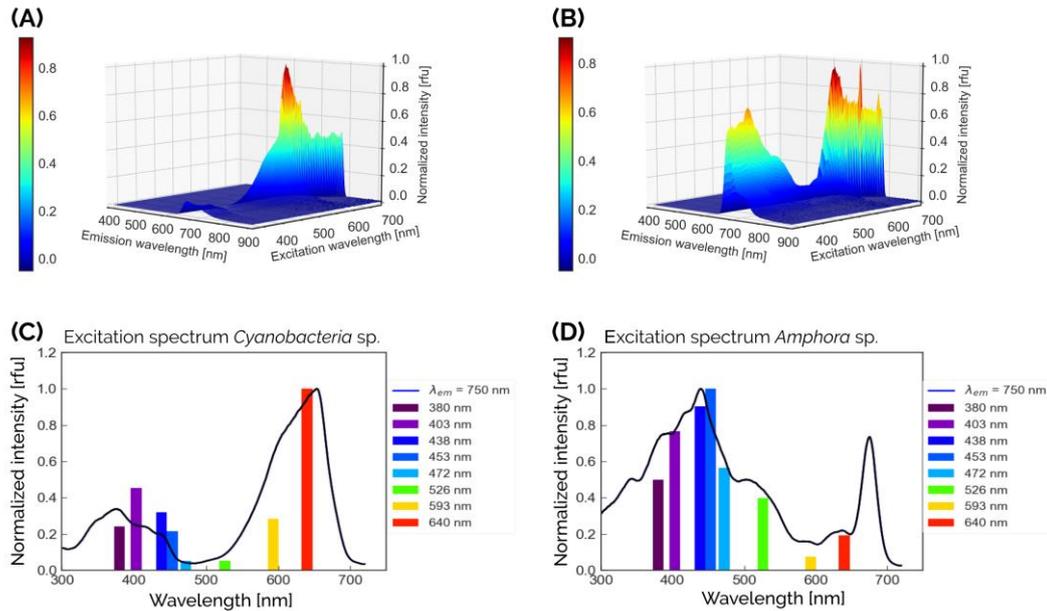
Figure S4-4 Top: Two FEEMs of a cyanobacteria (**A**) and a diatom (**B**) sample illustrating the spectral differences of different algal groups according to their distinct pigment composition of their photosystems. In the FEEMs, the intensity is normalized to the maximal fluorescence intensity. Peaks above 700 nm excitation are first order diffraction artifacts. **Bottom:** Comparison of the relative fluorescence patterns recorded with the ALPACA and the reference measurement of the spectrofluorometer for the cyanobacteria (*Cyanobacteria* sp.) (**C**) and the diatom (*Amphora* sp.) sample (**D**). Apart from LED 593 nm, which in general exhibits quite low signal intensities, the relative fluorescence pattern coincides with the reference spectra.

4.6.7. Statistical Principle of the Linear Discriminant Analysis

The linear discriminant analysis is a supervised pattern recognition analysis targeting the assignment of an unknown object to the correct group with respect to several spectral variables. The discriminant analysis seeks the optimization of Fisher's criterion: the separation of objects belonging to different groups while the group variability should be small as possible.

$$J(\omega) = \frac{\omega^T \cdot S_B \cdot \omega}{\omega^T \cdot S_W \cdot \omega} \qquad \text{with } w = \sum_{i=1}^{d} \alpha_i \cdot \varphi_i \qquad \text{S 4.2}$$

First, the pattern recognition technique determines the structure or characteristic properties of various training objects of known allocation. In our approach, the pattern recognition bases on the contribution of eight different excitation wavelengths to the fluorescence intensity emitted from the photosystems within the phytoplankton. These contributions are combined as weighed linear combinations (discriminating variables), which is the sum of all bars in the histogram ('fluorescence pattern') of Figure S4-5, multiplied by a pre-factor.

Afterwards, an unknown sample can be classified with regard to this training set (Figure S4-5, score plot). Consequently, the separation effectivity improves with the number of training objects and the

quality of characterization. For our investigations, we used 44 phytoplankton species as training objects belonging to nine different classes and 30 different algal orders:

Spirulinales (1), *Synechococcales* (1), *Bonnemaisoniales* (1), *Glaucosphaerales* (1), *Gonyaulacales* (1), *Gymnodiniales* (3), *Dinotrichales* (1), *Prorocentrales* (1), *Peridiniales* (2), *Pyrenomonadales* (4), *Isochrysidales* (2), *Prymnesiales* (4), *Pavlovales* (1), *Coccolithales* (1), *Sarcinochrysidales* (1), *Chattonellales* (1), *Eustigmatales* (1), *Chromulinales* (1), *Stephanodiscales* (1), *Cymatosirales* (1), *Thalassiosirales* (2), *Bacillariales* (1), *Chaetocerotales* (1), *Chlamydomonadales* (1), *Mamiellales* (3), *Nephroselmidales* (1), *Oltmannsiellopsidales* (1), *Pyramimonadales* (1), *Chlorodendrales* (1), *Eutreptiacaeae* (2)

As the linear discriminant analysis separates, each times, two classes from another, we obtained 29 different discriminant functions. The coefficients of the factors of each discriminant function are weighting factors (loadings) indicating the importance of the variables for the discrimination.

The unknown sample is assigned to the training groups where the fluorescence pattern fits best, meaning the physical distance is minimal.

Further information about the mathematical approach, its performance and its significance level will be provided discussed in a pursuing publication which is already in preparation.

4.6.8. Software Interface
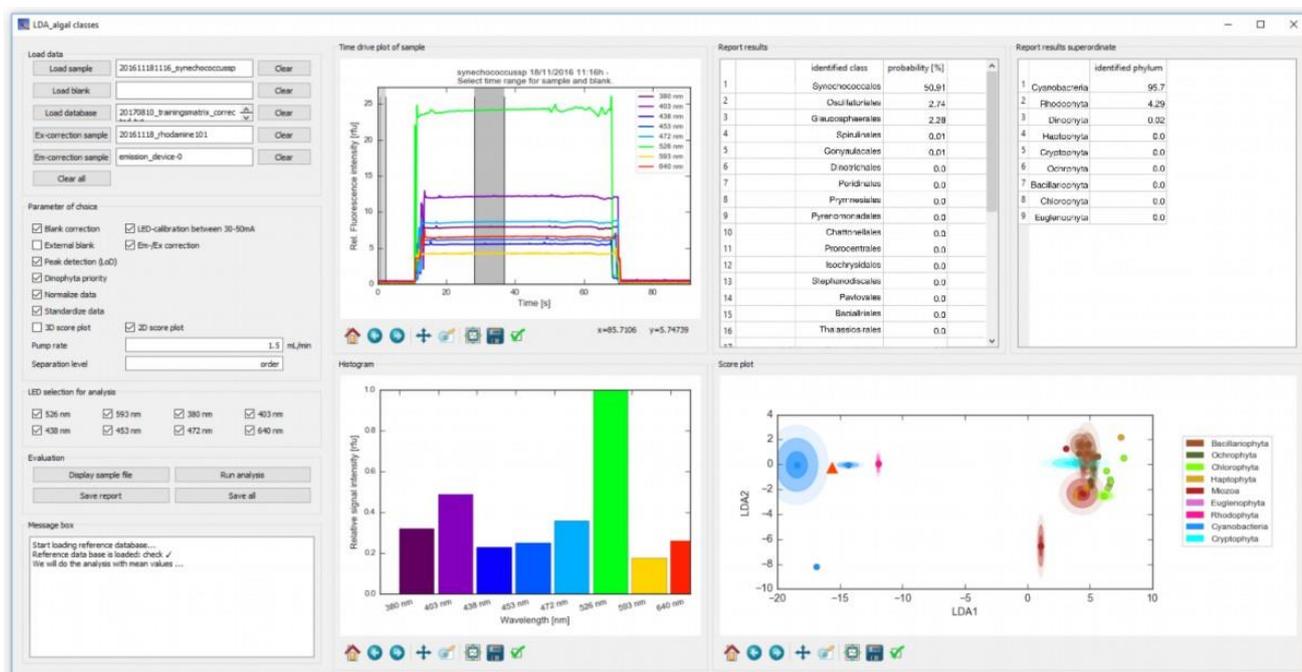
Figure S4-5: Software interface for data evaluation developed in Python – an open source programming language for user-friendly data evaluation. For demonstration, a highly-concentrated cyanobacteria suspension (*Synechococcus* sp.) is used. On the top, the fluorescence upon excitation at eight different excitation wavelengths is shown, whereas the bar plot below represents the fluorescence pattern of the cyanobacteria. On the right, the resulting score plot after applying the LDA is shown. The sample is visualized as orange triangle with respect to the training objects which are grouped as 'algal centroids'.

On the left panel of the interface, the user can select multiple parameters relevant for the evaluation (Figure S4-6). In particular, one can define, whether a blank correction or standardization strategy should be executed or not. Furthermore, the algorithm can be specified to focus on the distinct discrimination of dinoflagellates. Beside these process parameters, the user can choose which of the LEDs should be included in the data evaluation process. The possibility to exclude an LED can be used to investigate the influence of a specific excitation wavelength on the discrimination capability of the algorithm. Findings gained from these investigations can further reduce the amount of LEDs integrated into the algae module.

Figure S4-6: Selection of parameters relevant for data evaluation. Besides internal standardization strategies, the user can select which LED should be included for data evaluation.

On the right side of the software interface (Figure S4-5), the measurement signals recorded by the ALPACA and the evaluation results after applying the discriminant analysis are displayed. Besides the score plot of the sample and the trainings objects, a probability matrix of group membership is returned. This matrix describes the likelihood of group membership for the unknown sample after

analyzing spectral characteristics of the sample with regard to a well-defined trainings database. The discrimination is applied at the taxonomic 'order' and 'class level.

4.6.9. Linearity of the ALPACA and Limit of Detection (LOD)

Sulforhodamine 101 between $0.03-7.4\,\mu$M was used to test the linearity of each excitation channel of the ALPACA. Its spectral properties are shown in Figure S4-7. The fluorescence dye was excited with a blue 472 nm - LED for better comparability between the excitation channels. The emission was recorded on the silicon-PIN-photodiode either above the blocking range $\lambda_s = 650$ nm, if the detector was covered with a RG665 glass filter, or above the blocking range $\lambda_s = 710$ nm, if covered with a RG9 glass filter. The calibration curve for each measurement channel and calculated coefficients of determination are shown in Figure S4-8.
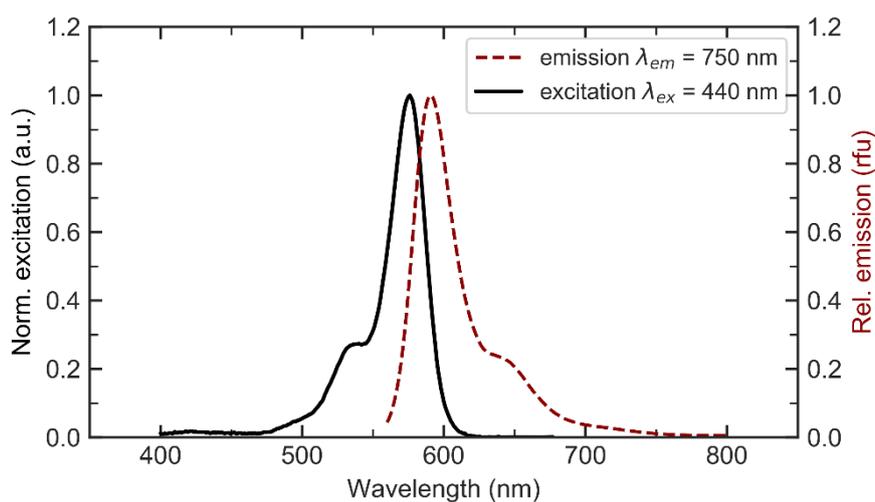
Figure S4-7: Spectral properties of sulforhodamine 101. Excitation spectrum was recorded at 750 nm and emission spectrum was recorded upon excitation at 440 nm.
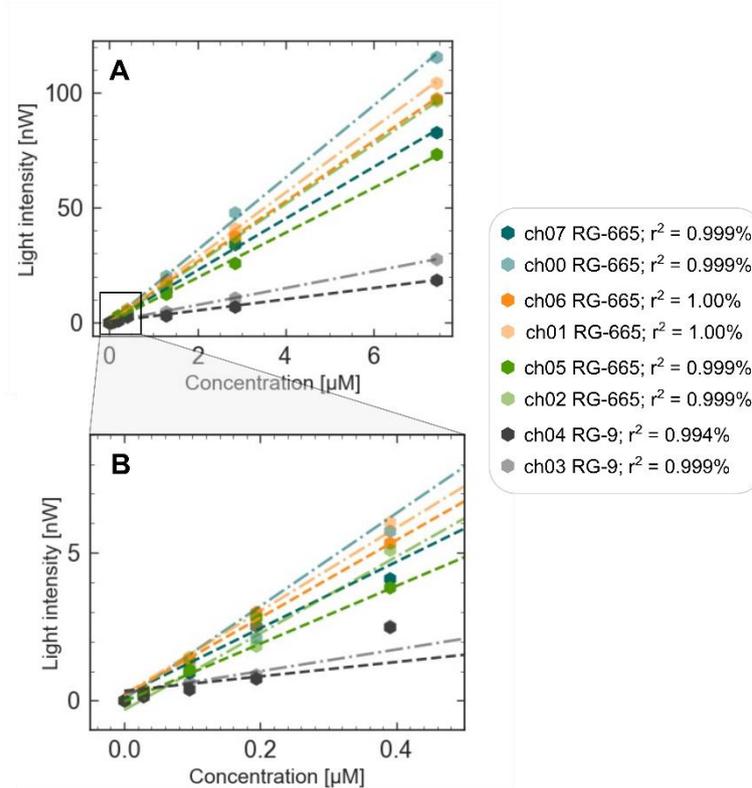
Figure S4-8: Linearity check and coefficient of determination for each measurement channel of the ALPACA using sulforhodamine 101 solution between 0.03–19.7 μM. (**A**) The fluorescence standard was excited with a 472 nm-LED at each excitation channel. The overall fluorescence emission was recorded either above 650 nm (blocking range), if the photodiode was covered with an RG-665 emission filter, or above 710 nm, when covered with an RG-9 emission filter. (**B**) Zoom-in to lower concentrations between 0.03–0.4 μM.
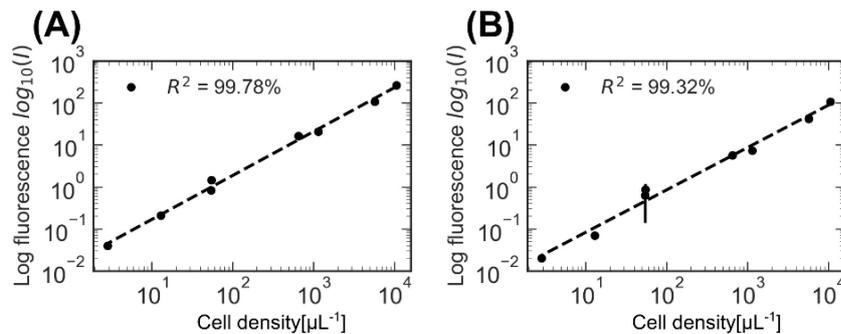
Figure S4-9: Correlation between average cell density of the diatom *Amphora* sp. (cell dimensions (width x length): 6 x 21 μm²) and recorded fluorescence intensity upon excitation at 453 nm (**A**) and 526 nm (**B**) in a double-log depiction including the standard deviation (n=3). The biomass varied between 1.0 cell per microliter and 10.5×10^3 cells per microliter.
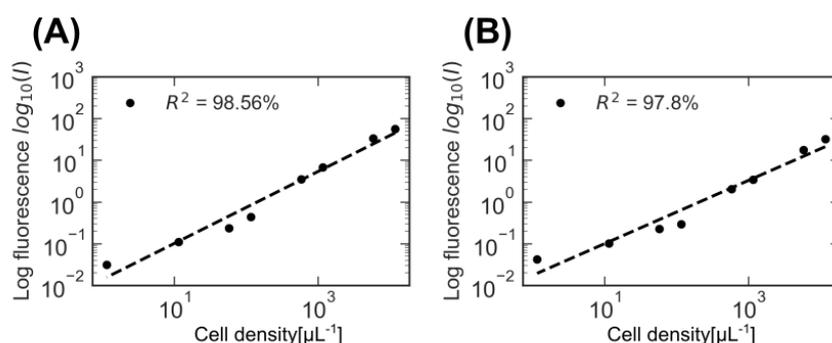
Figure S4-10: Correlation between average cell density of *Cyanobacteria* sp. (cell dimensions (width x length): 4 x 111 μm²) and recorded fluorescence intensity upon excitation at 453 nm (**A**) and 526 nm (**B**) in a double-log depiction including the standard deviation (n=3). The biomass varied between 1.0 cell per microliter and 11.6 × 10³ cells per microliter.
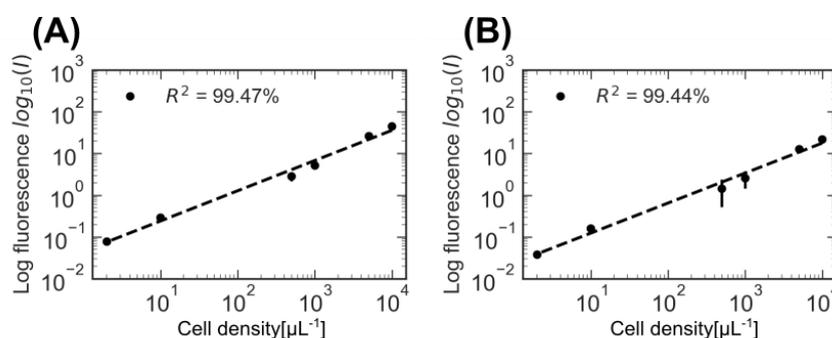
Figure S4-11: Correlation of average cell density of the cryptophyta *Hemiselmis cf rufescens*. (cell dimensions (width x length): 8–7 μm²) and recorded fluorescence intensity upon excitation at 453 nm (**A**) and 526 nm (**B**) in a double-log depiction including the standard deviation (n=3). The biomass varied between 2.0 cell per microliter and 10.0 × 10³ cells per microliter.

4.6.10. Limit of Detection Experiment and Single Cell Detection of the Diatom *Amphora* sp.

In the course of correlation between biomass and relative fluorescence intensity, the detection limit of the ALPACA was investigated. Figure 4-3 in the main part illustrates how the fluorescence signal changes with biomass in the measurement chamber. While the signal is dense and compressed with a high average signal level at higher biomass concentrations, the signal decreases and gets resolved at lower cell concentrations. At lower biomass concentrations, individual signal spikes correspond to individual cell events passing through the measurement chamber.

Even after a dilution to 0.2 cells per microliter, it was possible to record signal spikes corresponding to 1 cell per measurement volume. In the first seconds, the standard deviation of the background signals *I(n)* of the blank solution is detected. *I(s)* denotes the absolute, e.g. blank corrected signal intensity, when an algal cell was passed through the measurement chamber. The quotient of

absolute signal intensity *I(s)* and background signal *I(n)* determines the signal-to-background ratio and is calculated to ~ 3. This ratio increase with increased biovolume in the measurement chamber.
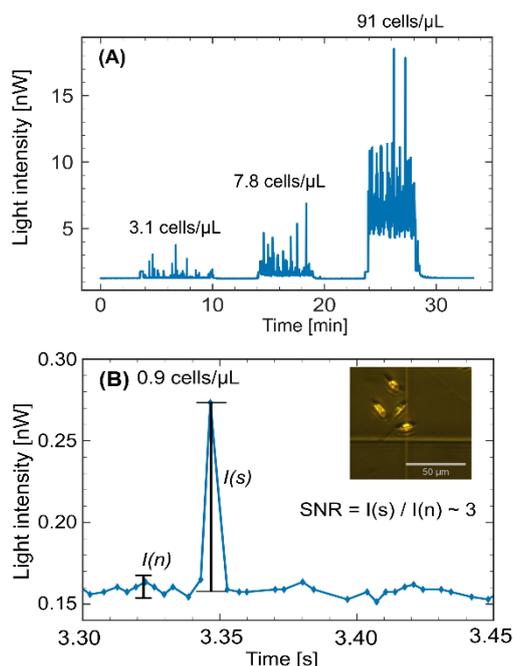
Figure S4-12: Variations of the fluorescence signal of the diatom *Amphora* sp. depending on the biomass between 3.1–91 cells per microliter upon excitation at 453 nm **(A).** When the biomass in the measurement chamber is high, the average signal level is heightened, but gets resolved at lower cell densities (3.1 cells per microliter). When the cell density is even lower, i.e. 0.9 cells per microliter, individual signal spikes correspond to individual cell events passing through the measurement channel allowing then a single cell detection and cell counting **(B).** The slight increase of background signal due to multiple scattering effects within the capillary cannot be completely prevented.

4.6.11. Statistical Classification and Quality Criteria for all Algal Groups

Table S4-4 Quality criteria for the evaluation of 18 algal samples belonging to eight different algal phyla using the multivariate discriminant analysis (LDA).[a]

algal group	number of samples	sensitivity (TPR) (%)	miss rate (FNR) (%)	specificity (TNR) (%)	accuracy (ACC) (%)
Cyanobacteria plus Rhodophyta	4	100.0	0.0	100.0	100.0
Dinophyt	3	100.0	0.0	86.7	88.2
Cyanobacteria	2	50.0	50.0	100.0	94.1
Rhodophyta	2	100.0	0.0	93.8	94.1
Haptophtyta	2	0.00	100.0	100.0	88.2
Bacillariophyta	3	33.3	67.7	92.9	82.4
Chlorophyta	4	100.0	0.0	100.0	100.0
Ochorphyta	1	100.0	0.0	100.0	100.0
Cryptophyta	1	0.0	100.0	100.0	94.1

[a] Acronyms: **TPR:** True-positive rate, **FNR:** False-negative rate, **TNR:** True-negative rate, **ACC:** accuracy

4.6.12. Sampling Site for Long-Term Field Campaign
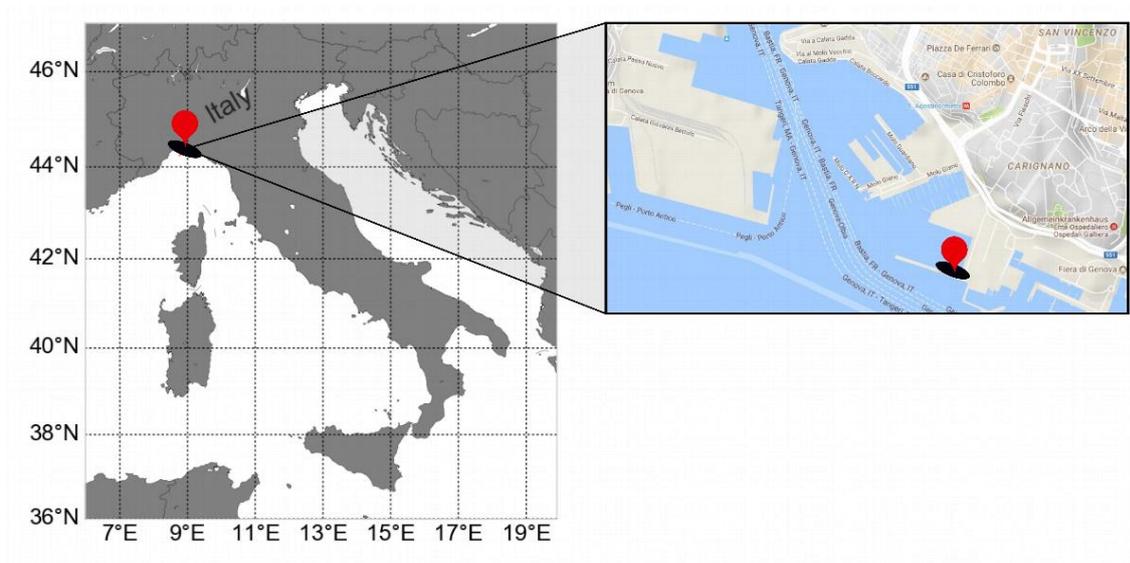
Figure S4-13: Sampling station for the long-term field campaign in the harbor of Genoa (Italy) as overview (**left**) and zoom in (**right**). The buoy was deployed at the entrance of the harbor.

5. Spectral Characterization of Eight Marine Phytoplankton Phyla and Assessing a Pigment-Based Taxonomic Discriminant Analysis for the in-Situ Classification of Phytoplankton Blooms
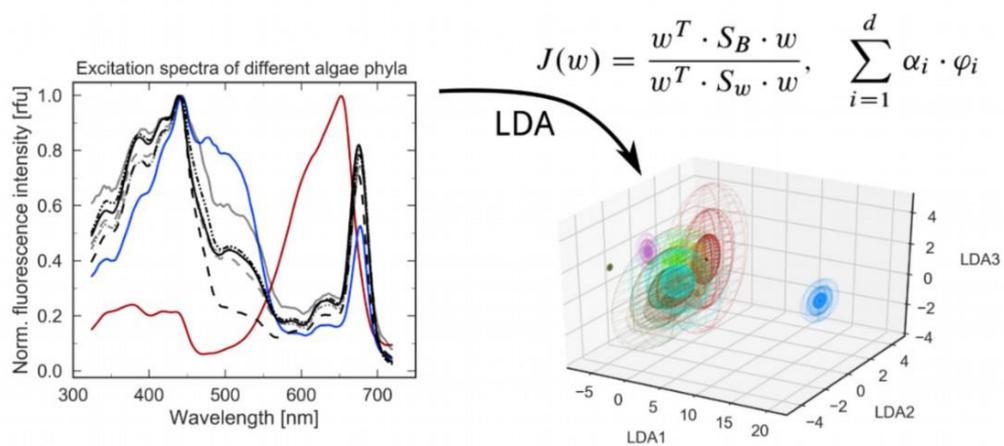
5.1. Preface for the Manuscript

This publication is dedicated to the theoretical approach of the ALPACA and represents its evaluation strategy based on a multivariate classification algorithm in order to emphasizes the applicability of pigment-based chemotaxonomic approaches in marine monitoring programs. One of these approaches is Fisher's linear discriminant analysis. This approach examines the inherent pattern of predefined trainings objects, aiming at the separation of objects that belong to different groups, while the variance within one group ought to be minimal. Although these approaches are underestimated in marine science, the results of the classification algorithm are promising with an overall classification capability of 97.3%.

As the ALPACA was designed as early-warning system for potentially harmful algae, the algorithm ought to identify algae groups that are potentially able to produce biotoxins which can accumulate across the food web. Thus, we started an extensive study on spectral properties of different marine phytoplankton groups and identified their hazard potential for marine life and humans. This identification was done in cooperation with the French Research Institute for Exploitation of the Sea (IFREMER) and the Plant biology and Ecology Department of the University of the Basque Country (UPV/EHU) in Spain. Based on their investigation, we decided to focused on the reliable separation of cyanobacteria and dinoflagellates as ubiquitous and well-known toxin producing algae. The study of the spectral properties enables then the extraction of wavelength ranges which are most relevant for discrimination. For further details, please refer to the supplemental material of this publication (section 5.6) or to chapter 2.

Besides the theoretical background of the pigment-based chemotaxonomic approach, we validated the systems performance with a leave-one-out cross validation with 53 different marine phytoplankton species and studied the separation capability in mixed samples. As the pigment patterns are known to change due to a variety of reasons, among them, due to different light conditions, growth phases or due to nutrient availability, we studied the system performance under different light conditions and at different growth phases (see section 5.5.2). The results of the study were promising that an adequate and upgraded trainings database, in which these alterations are considered, can deal with the challenges.

Spectral Characterization of Eight Marine Phytoplankton Phyla and Assessing a Pigment-Based Taxonomic Discriminant Analysis for the in-Situ Classification of Phytoplankton Blooms

Authors: Silvia E. Zieger[1#], Sergio Seoane[2], Aitor Laza-Martínez[2], Anna Knaus[1], Günter Mistlberger[1] and Ingo Klimant[1]

[1] Optical Sensors Group, Institute of Analytical Chemistry and Food Chemistry, Graz University of Technology, Graz, Austria

[2] Plant biology and Ecology Department, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Spain

 # Corresponding author: zieger@tugraz.at

5.2. Abstract

Early stage identification of harmful algal blooms (HABs) has gained significance for marine monitoring systems over the years. Various approaches for in-situ classification have been developed. Among them, pigment-based taxonomic classification is one promising technique for in-situ characterization of bloom compositions, although it is yet underestimated in marine monitoring programs. To demonstrate the applicability and importance of this powerful approach for monitoring programs, we combined an ultra-low-cost and miniaturized multi-channel fluorometer with Fisher's linear discriminant analysis (LDA). This enables the real-time characterization of algal blooms at order level based on their spectral properties. The classification capability of the algorithm was examined with a leave-one-out cross validation of 53 different unialgal cultures conducted in terms of standard statistical measures and independent figures of merits. The separation capability of the linear discriminant analysis was further successfully examined in mixed algal suspensions. Besides this, the impact of the growing status on the classification capability was assessed. Further, we provide a comprehensive study of spectral features of eight different phytoplankton phyla including an extensive study of fluorescence

excitation spectra and marker pigments analyzed via HPLC. The analyzed phytoplankton species belong to the phyla of Cyanobacteria, Dinophyta (Dinoflagellates), Bacillariophyta (Diatoms), Haptophyta, Chlorophyta, Ochrophyta, Cryptophyta and Euglenophyta.

5.3. Introduction

The evident increase in intensity, frequency and geographic distribution of algal blooms widely spread all over the world has focused attention on the intrinsic physiological, ecological, and toxicological aspects of individual phytoplankton species involved.[3] Some algal blooms are restricted to local environments and well-studied in their composition and their effects on the local environment and human health. Others, however, are almost undescribed. Many harmful algal species expand to regions where they were previously unknown, becoming then bloom-forming in these areas.[143] In the same way, the potential risk of human disorders after uptake of contaminated water has become a major concern.[4] The increase in frequency and intensity raises therefore awareness to study algal distributions, dynamics and individual bloom compositions all over the world.

Conventional bloom identification and quantification methods rely mainly on morphological studies or on pigment-based assessments.[126,144–146] Morphological approaches focus on cellular properties and distinguish algal groups on the basis of their scattering properties, among other things. The standard approaches for the taxonomic identification and quantification based on cellular properties are microscopes and flow cytometers. Although they have proven to be efficient in distinguishing and quantifying biomass in-lab, they are often expensive and time-consuming.[10,22] In contrast to morphological approaches, pigment-based approaches using fluorometric measurements have proven to be a promising tool to examine bloom compositions in-situ.[127,128,147,148] The relative composition of phytoplankton pigments can provide valuable taxonomical information at class level. The evaluation of present or absent marker pigments can therefore be used to identify phytoplankton community compositions. Besides qualification, studying chlorophyll-a concentrations enables also the quantitative control of episodic alterations of marine life.[10,149] However, pigment-based approaches are still underestimated in marine monitoring programs. Owing to the lack of appropriate monitoring tools, we have developed a miniaturized, ultra-low-cost device which enables the taxonomic classification of algal assemblages. Its fabrication, operation and application in coastal environments was presented earlier[150], while the data evaluation procedure and software performance is presented here. This monitoring device, named ALPACA, examines the characteristic pigment pattern upon excitation with eight light-emitting diodes (LEDs) in the visible wavelength range (380–640 nm). Contribution of different (marker) pigments to the resulting chlorophyll fluorescence are then examined by applying a multivariate discriminant algorithm.

The analysis of choice is Fisher's linear discriminant analysis. This discriminant analysis is a supervised pattern recognition method, targeting the allocation of an unknown object to the correct class with respect to several spectral variables.[80] Therefore, the pattern recognition technique requires a pool of well characterized and labeled training objects. To enhance the allocation and identification effectivity, a comprehensive characterization of spectral properties of the training objects is one main aspect of the pattern recognition technique. Another aspect is the extraction of key features which are significant for the separation of different phytoplankton. Furthermore, separation effectivity improves with the number of training objects.[51,148,151,152]

After publishing the fabrication and promising application of the ALPACA during various field tests[150], the data evaluation method for in-situ discrimination of phytoplankton groups is presented here. The discrimination method bases on distinctive pigment pattern and therefore this study contains three different objectives: First, we characterized the spectral properties of the phytoplankton groups and extracted their key differences using fluorescence excitation and emission spectra using HPLC. Second, we present the multivariate discriminant analysis with optimized separation capability. Thereby the focus of our studies was the reliable differentiation of cyanobacteria and dinoflagellates from co-occurring algal groups. Finally, we tested the system performance of the pattern recognition algorithm under different growing conditions regarding two factors: (a) during different growth phases and (b) under different scalar irradiances.

5.4. Material and Methods

5.4.1. Phytoplankton Cultures

A total of 53 phytoplankton strains were obtained from the Culture Collection of the Plant biology and Ecology Department from the University of the Basque Country (UPV/EHU) (Spain). The strains belong to 8 different phyla and 27 orders. Each strain was cultured in conical flasks with f/2 growth medium at a salinity of 33 and with addition of selenium and silicate as appropriate.[131,132] The cultivating temperature was 17 °C and the irradiances with a cool white fluorescent tube were 20, 100 or 200 μmol / (s·m^2). The light:dark cycle during cultivation was 10:14 h. Further detail of the taxonomic classification of the phytoplankton samples is given in Table S5-1 in the supplemental material.

5.4.2. Sample Preparation

Before each measurement, each phytoplankton sample was analyzed under the light-inverted microscope (Leica Microsystems GmbH, Wetzlar, Germany) equipped with a DS-U2 camera

(Nikon, Instruments Europe, Badhoevedorp, the Netherlands) to ensure that the cultures were healthy and without contamination. The cell density was determined following the method described by Andersen.[136]

5.4.3. HPLC Analysis

Pigment extraction, preparation and analysis were performed under low light following the method described by Zapata et al.[77] with a modification in solvent A, explained in Seoane et al.[153] Variable volumes of cultures were filtered onto Whatman GF/F glass fibre filters (25 mm diameter, Whatman International Ltd, Maidstone, UK) under vacuum. The filters were stored immediately at -80 °C until used, but less than 15 days after filtration. For extraction under low light, frozen filters were extracted in screw capped tubes with 3 mL 90% acetone using a glass rod for grinding. Extracts were then filtered through 0.2 μm syringe filters to remove cell and filter debris. An aliquot of acetone extract (0.5 mL) was mixed with 0.2 mL of water and the resulting solution was immediately injected. The chromatography equipment run with Millenium32 software. It was an Alliance HPLC System consisting of a 2695 separations module, a 996 diode-array detector (1.2 nm optical resolution), and a 474 fluorescence detector, all by Waters (Milford, MA, USA). The column was a Waters Symmetry C8, 150 ×l4.6 mm, 3.5 μm particle size, 100 A pore size. Two eluents were used, eluent A (a mixture of methanol: acetonitrile: aqueous pyridine (0.25 M) (50:25:25; v:v:v)) and Eluent B (a mixture of methanol:acetonitrile:acetone (20:60:20; v:v:v)), and the solvent mixture gradient profile was the following, time, % A solvent: 0 min, 100%; 22 min, 60%; 28 min, 5%; 38 min, 5%; 42 min, 100%. The temperature of the column was 30 °C and the flow rate 1 mL min^{-1}. Pigments were identified by their retention times and absorbance spectra. The absorbance chromatograms were extracted at 440 nm. Retention times were compared with those of pure standards obtained commercially from DHI (Hoersholm, Denmark) and those reported in Jeffrey[154] and Zapata et al.[77] HPLC calibration was performed using chlorophyll (a, b and c_2) and carotenoids (fucoxanthin, lutein, zeaxanthin, peridinin and alloxanthin) standards obtained commercially from DHI. The molar extinction coefficients obtained from Jeffrey[154] were used for the quantification of pigments not calibrated with commercial standards.

5.4.4. Fluorescence Spectra

5.4.4.1. Excitation and Emission Spectra of the Full Spectral Range

Fluorescence emission and excitation spectra were acquired for each algal sample with a fluorescence spectrofluorometer (Fluorolog–3, Horiba Jobin-Yvon, France) at room temperature. According to Poryvkina et al.[38] and Callieri et al.[155], the dark-adapted samples were held in

suspension with glycerin to avoid sinking of the cells during measurement. For the excitation spectrum, the excitation wavelength was scanned between 300–730 nm (slit width 7 nm, increment 1 nm) and the emission wavelength was set to 750 nm (slit width 14 nm) in order to reveal spectral properties of chlorophyll-a and auxiliary pigments in the light-harvesting complex. The detector was additionally covered with an RG-665 longpass filter (bk Interferenzoptik, Germany) to avoid spectral artifacts. To record the emission spectrum, the excitation wavelength was set to 595 nm (slit width 7 nm, increment 1 nm) in the case of cyanobacteria to trigger their major group of light-harvesting pigments, the phycobiliproteins. The emission wavelength was scanned between 630–800 nm (slit width 14 nm). In all other cases the excitation wavelength was set to 440 nm and the emission wavelength was then scanned between 550–800 nm with similar slid widths as before. Both spectra were corrected internally to compensate the wavelength dependence of the optical components[45]. Additionally, each spectrum was compared to its maximal signal intensity leading to the relative excitation or emission spectrum.

5.4.4.2. Relative Fluorescence Pattern at Distinct Excitation Wavelengths

Discontinuous fluorescence spectra were acquired for each algal sample with the miniaturized, low-cost ALPACA, which has been developed in the frame of the FP7-SCHeMA-project (2015) and which was presented earlier.[5,150] The setup of the ALPACA, an acronym for Advanced Luminescence-based Phytoplankton Analysis and Classification Appliance, is shown in Figure 5-1D. The detection unit of the ALPACA features eight excitation channels – four at each side – and four emission channels aligned at right angles. During operation, up to eight different LEDs (380 nm, 403 nm, 438 nm, 453 nm, 472 nm, 526 nm, 593 nm, 640 nm, selected for spectral and technical reasons) are used to excite the aqueous phytoplankton solution, which is pumped through the capillary in the detection unit. The spectral properties of the LEDs are shown in Figure 5-1C. Surplus energy, which cannot be used for photosynthesis, is re-emitted by the phytoplankton in the capillary and recorded by the photodiodes as accumulated light intensity in the range between 650-800 nm. The recorded emission window is defined by two longpass filters as it is described in previous publications[150]. The fluorescence intensity emitted from the phytoplankton depends on the efficiency of excitation at a specific excitation wavelength, or respectively the chosen LED, and varies with the pigment composition in the photosystem of the phytoplankton. Consequently, the ALPACA provides a relative fluorescence spectrum at distinct excitation wavelengths, leading to a bar plot, that represents a discontinuous description of excitation properties of the sample. The comparison between the excitation spectrum and the bar plot gathered from the ALPAC is shown in the supporting information (see page 109). The fluorescence intensity upon excitation, which is recorded by the ALPACA, is analyzed by the multivariate discriminant analysis for classification of different phytoplankton groups.

5.4.5. Fisher's Linear Discriminant Analysis for Determination of Class Membership

5.4.5.1. Theoretical Background

Fisher's linear discriminant analysis is a statistical, supervised pattern recognition method. It uses a coordinate transformation and projection into a subspace to enhance the separation of variables with respect to their features, in particular their taxonomic membership. As pattern recognition algorithm, the discriminant analysis requires a set of training objects to allocate a new object of unknown group membership.[80,156] Fisher's linear discriminant analysis seeks linear combinations of measurement variables that optimizes the separation of objects from different groups. According to the Fisher's criterion, the separation of K groups is optimal, if the variability within one group is minimized while the variability between different groups is maximized

$$J(w) = \frac{w^T \cdot S_B \cdot w}{w^T \cdot S_W \cdot w} \quad \text{and} \quad w = \sum_{i=1}^{d} \alpha_i \cdot \varphi_i \qquad \textbf{5.1}$$

with $w = (w_1, w_2, \ldots w_d)^T$ is the vector of discriminants representing a weighted linear combination of measurement features φ_i, S_w and S_B are the within- and between-scatter or covariance matrices of the data matrix. The solution to the separation problem can be accomplished using Lagrange multipliers $\mathcal{L} = w^T \cdot S_B \cdot w - \lambda[(w^T \cdot S_w \cdot w) - 1]$. This Lagrangian can be transformed to an optimization or generalized eigenvalue problem $S_B \cdot w = \lambda \cdot S_w \cdot w$, assuming the inverse of S_w exists. The resulting eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\min(K-1,d)}$ and eigenvectors $w_1 \geq w_2 \geq \cdots \geq w_{\min(K-1,d)}$ generate corresponding discriminant vectors $z_1, z_2, \ldots z_{\min(K-1,d)}$ where the first eigenvalue λ_1 and its corresponding discriminant vector z_1 gives the largest measure of Fisher's criterion. The discriminant vectors of the training objects span a sub-dimensional space, where the relationship among different groups or centroids can be studied as it is shown in Figure 5-2.[134]

To predict the group membership of an unknown sample, a decision function based on the Mahalanobis distance ∂ between the sample and each group centroid is calculated

$$\partial(x, k_i) = \sqrt{(x - k_i)^T \Sigma^{-1} (x - k_i)} \qquad \textbf{5.2}$$

With Σ is the covariance matrix, x is the score vector of the sample after transformation and projection into the subspace and k_i is the median of the group centroid in the subspace. The distance is combined with a multi-dimensional Gaussian distribution which determines the probability of group membership $\wp$ for each group in the training matrix:

$$\wp = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma}} e^{(-t)}, \quad with \quad t = \frac{(\partial - \mu)^2}{2 \cdot \sigma^2} \qquad \textbf{5.3}$$

The normal density function can be used in its reduced form, as each of the objects were normalized and standardized ($\sigma = 1$, $\mu = 0$) in advance.

For classification and separation of the phytoplankton groups, in particular cyanobacteria and dinoflagellates from co-occurring algae, Fisher's linear discriminant analysis was applied to the measurement data gained from the ALPACA. Herein, the fluorescence intensities of a sample at a distinct excitation wavelength are the measurement features φ_i of the linear combinations in equation 5.1. For further information about the input variables are provided in the supporting information. These linear combinations w consist therefore of eight measurement features for each phytoplankton group.

All data were processed using a graphical interface written in Python.

5.4.5.2. Spectral Characterization and Feature Extraction

Feature extraction is the key to face recognition to reduce dimensionality of parameters, so that the extracted features are as representative as possible. In pigment-based analyses, recognition and classification, both base on spectral characteristics within and between phytoplankton groups. To examine the spectral characteristics of different phytoplankton groups used as training objects for the discriminant analysis, we characterized 53 phytoplankton species belonging to eight different phyla threefold: First, we recorded the fluorescence excitation and emission spectra using the fluorescence spectrophotometer to get an initial hint of spectral similarities and differences within one class and between different groups. Second, we investigated the marker pigments, which are theoretically valuable for distinct identification of one phytoplankton group, by HPLC analysis. In addition, we investigated the typical excitation wavelengths of these marker pigments in the photosystem of the phytoplankton. Following the spectral characterization, we selected the excitation ranges, which are considered most reliable for a pigment-based chemotaxonomy.

Figure 5-1A-B summarizes typical excitation and emission spectra of the eight investigated phytoplankton phyla. While the emission properties of all phytoplankton phyla are quite similar in shape and absolute emission wavelength (Figure 5-1B), the excitation properties vary between different phytoplankton groups (Figure 5-1A). Classification methodologies therefore focus on excitation properties of phytoplankton species instead of their emission properties. Cyanobacteria mainly absorb between 600−670 nm, where other algal groups absorb only to a small extend. These algal groups mainly absorb at shorter wavelengths between 400−500 nm. The distinct shape of the algal excitation spectrum, due to the relative pigment composition of their photosystems, was investigated via HPLC analysis. An extract of the most important lipophilic marker pigments examined by HPLC analysis is shown in Table 5-1. The pigments are presented in terms of marker pigment : chlorophyll-a ratios. The values are averaged along one algal order. Detailed information about the relative pigment composition examined for different phytoplankton phyla are given in the supplemental material (Table S5-2 to Table S5-10).
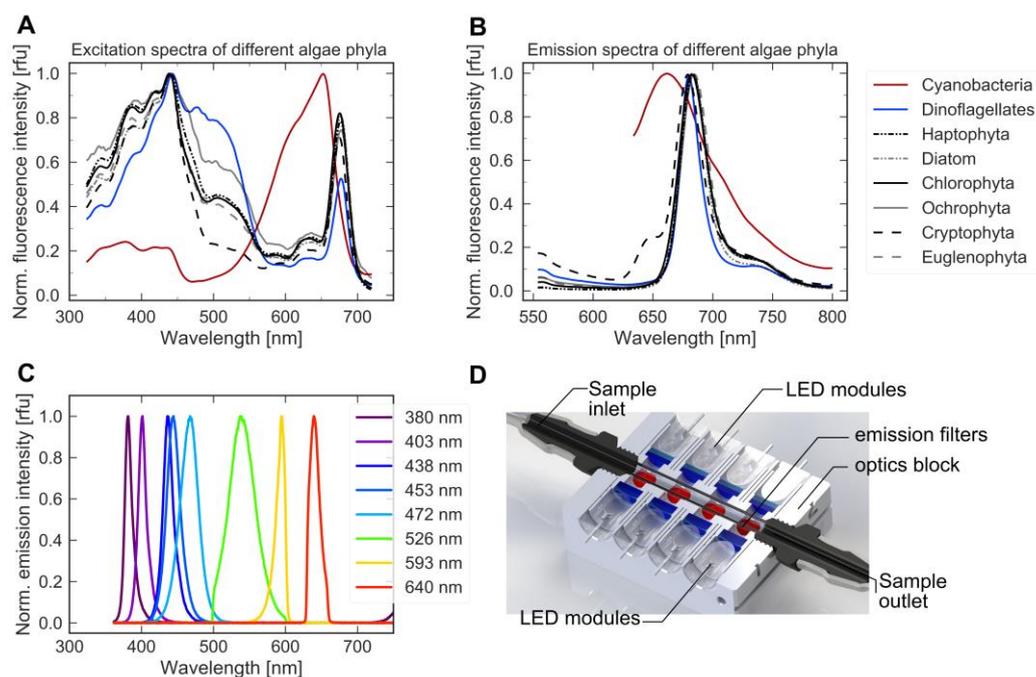
Figure 5-1: Excitation **(A)** and emission **(B)** spectra of eight algae species belonging to different algal classes to demonstrate spectral characteristics and differences between these algal groups. Algae species examined are *Synechococcus* sp. (Cyanobacteria, marked in red), *Alexandrium minutum* (Dinophyta, marked in blue), *Emiliania huxleyi* (Haptophyta), *Thalassiosira weissflogii* (Bacillariophyta), *Chlamydomonas coccoides* (Chlorophyta), *Heterosigma akashiwo* (Ochrophyta), *Chroomonas collegionis* (Cryptophyta) and *Eutreptiella gymnastica* (Euglenophyta). **(C)** Eight LEDs with narrow emission angle were used in ALPACA for distinct excitation of relevant marker pigments. **(D)** Schematic cross-section of the optical, electronics and fluidics setup of the ALPACA presented in top view.

Table 5-1 Marker pigments identified by HPLC expressed as pigment : chlorophyll-*a* ratio (in mol averaged per order).[a]

phylum / class	per	19HF	19BF	allo	dino	fuco	lut	zea	chloro-*b*	chloro-*a*
cyanobacteria	--	--	--	--	--	--	--	1.39	--	1.00
dinoflagellates	0.89	0.26	0.10	--	0.06	0.56	0.11	0.02	0.54	1.00
type 1	0.89	--	--	--	0.06	--	--	--	--	1.00
type 2	--	0.26	0.10	--	--	0.40	--	--	--	1.00
type 3	--	--	--	--	--	0.75	--	--	--	1.00
type 4	--	--	--	--	--	--	0.11	0.02	0.54	1.00
diatom	--	--	--	--	--	0.59	--	--	--	1.00
haptophytes	--	0.41	0.13	--	--	0.47	--	--	--	1.00
ochrophytes	--	--	--	--	--	0.60	--	0.25	--	1.00
chlorophytes	--	--	--	--	--	--	0.09	0.02	0.61	1.00
euglenophytes	--	--	--	--	--	--	--	--	0.59	1.00
cryptophytes	--	--	--	0.52	--	--	--	--	--	1.00

[a] The characterization is done at class level. Dinoflagellate types according to plastid origin: 1, original dinoflagellate plastid; 2, haptophyte plastid; 3, diatom plastid; 4, chlorophyte plastid.. *Abbreviations:* ***per:*** *peridinin,* **19HF:** *19'-hexanoyloxy-fucoxanthin,* **19BF:** *19'-butanoyloxy-fucoxanthin,* ***allo:*** *alloxanthin,* ***dino:*** *dinoxanthin,* ***fuco:*** *fucoxanthin,* ***lut:*** *lutein,* ***zea:*** *zeaxanthin,* ***chloro-b:*** *chlorophyll-b,* ***chloro-a:*** *chlorophyll-a*

While certain phytoplankton classes can clearly be identified by the presence or absence of one marker pigment, other classes can only be identified after combining information of several marker pigments and their corresponding absorption wavelength ranges (Table 5-1). Thus, information from both, excitation spectra and pigment analysis were combined and potentially relevant wavelength ranges were selected. Thus, we were able to define eight excitation wavelengths as key features which are considered most reliable for the discriminant algorithm. Then, appropriate LEDs with narrow emission angle were selected for the ALPACA taking into account the optical window specified by the longpass emission filters. Besides their valuable contribution for the discrimination, their absolute intensity as well as their long-term stability were considered. The selected LEDs are shown in Figure 5-1C–D: One LED is used to excite pigments of the phytoplankton sample in the red wavelength range (640 nm) and two LEDs excite in the green to orange wavelength range (526 nm and 593 nm), although this wavelength range is undermanned for technical reasons. The remaining LEDs excite marker pigments at shorter wavelengths (380 nm, 403 nm, 438 nm, 453 nm and 472 nm).[150] The relative fluorescence intensities upon excitation at these specific wavelength ranges are used as measurement features φ_i for the linear combinations in equation 5.1 when applying the linear discriminant analysis.

5.4.5.3. Data Evaluation - Score Plots and Class Assignment

To establish the training database, the relative fluorescence patterns of 53 different phytoplankton species, belonging to 27 different orders and 8 phyla, were recorded using the ALPACA and analyzed by LDA. The training objects are labeled and classified by their group membership at a certain taxonomic level. The number of discriminant functions depends on the number of groups included in the training database. LDA compute the resulting scores for each of the 53 phytoplankton strains and then, the scores are separated by groups at the order level. The generated groups can be graphically presented in the score plot, shown in Figure 5-2. Within the score plot, the relationship and spectral similarities of different groups can be inspected and all training objects, belonging to the same group, are combined as one group centroid. The first, second and third standard deviation of each dimension define the three contours of the centroids. The standard deviations were calculated with regard to the group median. An overview of the standard deviations for each phytoplankton class is given in Table S5-12 of the supplemental material. The overall objective of this approach is to separate as much phytoplankton groups as possible, with special focus on the reliable separation and classification of cyanobacteria and dinoflagellates as well-known toxin producing algal phyla.
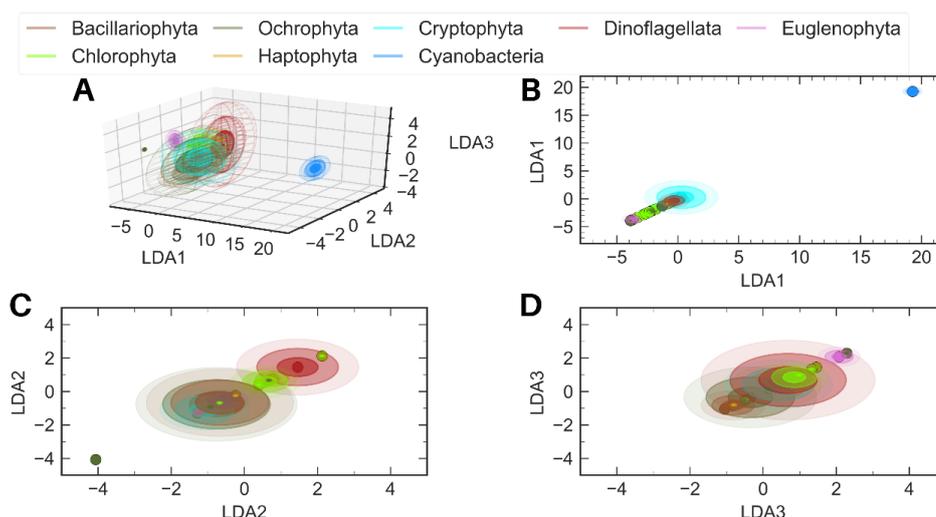
Figure 5-2: Score plot of training data (**A**) and two-dimensional projections (**B–D**). The two-dimensional projections illustrate which axis contains information on the original data variance enabling the separation of the phytoplankton groups. The centroid boundaries correspond to the first, second and third standard deviation of the algal groups. Separation of training objects is performed at class level for demonstration purpose.

The sub-dimensional score space is spanned by a definite number of projection axes. The first eigenvalue λ_1 and its corresponding discriminant vector w_1 generate the first projection axis, called *LDA1*, where the corresponding data variation is maximal in its projection direction (Fisher's criterion). This new projection axis explains the most of the original data scatter (see Table S5-13 in the supplemental material). The characteristic feature of this projection axis is the spectral difference between cyanobacteria and other algae, meaning that cyanobacteria are separated clearly along this projection direction (see Figure 5-2B). Further, the first projection axis promotes the separation of cryptophytes from euglenophytes and chlorophytes. The next projection axis, *LDA2*, reflects then a further part of the remaining differences between the groups. *LDA2* is useful to separate dinoflagellates from cryptophytes, haptophytes, ochrophytes and euglenophytes. The two axes span a 2D subspace in which the majority of the original data variation is explained. However, the separation along only two projection axes is often insufficient. By adding the third projection axis, the discriminant analysis is also able to separate dinoflagellates from diatoms and chlorophytes with only a narrow overlap area.

A set screw for our mathematical algorithm is the taxonomic level at which the algal groups are defined. Comparing the separation at different taxonomic levels, the explained variance accumulated over three projection axes is maximal for a separation at order level. In this case, the 3D subspace explains 99.92% of the original data scatter and with 4 discriminant vectors the whole data scatter is explained (see Table S5-13 in the supplemental material). The 3D score space is mainly used for visualization and graphical inspection. However, the 4D space is used for enhanced separation of the phytoplankton groups. Further information about the operational features for the discriminant algorithm are given in the supplemental material. To assign the class membership of the unknown sample, the Mahalanobis distance ∂ between the scores of the sample and each

median of the group centroid is calculated according to equation 5.7. The smaller the distance, the higher the spectral similarity for the unknown sample and thus, the higher the probability of group membership. Finally, a cut-off level was introduced to optimize the correct group assignment of cyanobacteria and dinoflagellates. During validation, we found out that the algorithm works optimal, if the difference of calculated probabilities for the first two algal groups is equal or higher than 4% (statistically significant; see the Excel sheet for further detail).

5.4.6. Validation and System Performance Assessment

5.4.6.1. Leave-One-Out Cross Validation

The system performance of LDA was investigated following the method described in Casson et al..[157] For the leave-one-out cross validation, one sample was excluded from the training database and its group membership was then assigned according to the remaining training objects. Following the evaluation of each phytoplankton sample, the true positives (T_P), true negatives (T_N), false positives (F_P) and false negatives (F_N) rates of the discriminant analysis were determined. The performance was then evaluated according to Batista[158] and Provost[159] in terms of sensitivity, specificity, selectivity and accuracy which are defined as the follows:

Accuracy or confidence level is a measure for correctly classified events compared to all predictions made

$$\text{accuracy (\%)} = \frac{T_p + T_n}{T_p + F_p + T_n + F_n} \qquad \textbf{5.4}$$

Selectivity is a measure for correctly identified events compared to all events classified to one group

$$\text{selectivity (\%)} = \frac{T_p}{T_p + F_p} \qquad \textbf{5.5}$$

Sensitivity is a measure for true-positive events for one specific group

$$\text{sensitivity (\%)} = \frac{T_p}{T_p + F_n} \qquad \textbf{5.6}$$

Specificity is a measure for true-negative events for one specific group

$$\text{specificity (\%)} = \frac{T_n}{T_n + F_p} \qquad \textbf{5.7}$$

Error Rate or prediction error is an overall measure of misclassified events

$$\text{error rate (\%)} = \frac{F_n + F_p}{T_p + F_n + F_p + T_n} \qquad \textbf{5.8}$$

The leave-one-out cross validation was conducted three times: First, the entire excitation spectrum recorded with the fluorescence spectrophotometer was analyzed with LDA. Second, the relative

fluorescence pattern was recorded with the ALPACA using all eight LEDs and third, the relative fluorescence pattern was recorded on the ALPACA using only four LEDs, whose absolute scales have the strongest impact on the discriminant vectors and who are not dependent (380 nm, 438 nm, 453 nm and 526 nm, see Table S5-14 and Table S5-15). All samples were cultivated under normal light conditions and were analyzed in the exponential growth phase.

5.4.6.2. Impact of the Light Intensity on the Relative Pigmentation along the growth Cycle

Alterations in the relative pigmentation under different growing conditions were examined in three different phytoplankton species: *Emiliania huxleyi* (*E. huxleyi)*, *Thalassiosira weissflogii* (*T. weissflogii)* and *Synechococcus* sp. as typical representatives of the major algal phyla Haptophyta, Bacillariophyta (Diatoms) and Cyanobacteria, respectively. After inoculation, each of the three samples was cultivated in conical flasks under low, medium and high light irradiance (20, 100 and 200 μmol / (s m²), respectively). The biomass was determined every day with the fluorescence spectrometer. In each growth phase, the relative pigmentation of the phytoplankton samples was analyzed by HPLC analysis and the relative fluorescence pattern was recorded by the ALPACA. Following the spectral characterization, the group membership was then investigated with LDA. Detailed information on the relative fluorescence intensity over the growth phase and the classification results are shown in the supplemental material (Table S5-17 to Table S5-36).

5.4.6.3. Mixed Sample Analysis

To elucidate the capability of the multivariate discriminant algorithm to identify minor compounds in mixed samples, a prokaryotic cyanobacteria *Synechococcus* sp. and the eukaryotic diatom *T. weissflogii* were mixed in eleven different ratios of cell density. Both cultivated under 100 μmol / (s m²). Their mixing ratios are shown in Table 5-4. The relative fluorescence pattern of each mixture was recorded at eight excitation wavelengths with the ALPACA as described before. The probability of group membership was then predicted using the multivariate discriminant analysis.

5.5. Results and Discussion

After implementation and spectral characterization of the training samples required for the multivariate discriminant analysis, LDA was validated in terms of standard figures of merit and its separation effectivity under different growing conditions and in mixed samples. Detailed information about the spectral characterization and the feature extraction for LDA is provided in the supplemental material.

5.5.1. Validation and System Performance Assessment

While it is straightforward to differentiate cyanobacteria from other algae by comparing the contribution to the emitted fluorescence at two different excitation wavelengths, it is almost impossible to differentiate further algal groups without multivariate statistics. Due to the significant overlap in their spectral characteristics, the linear discriminant analysis was introduced to enable the differentiation of cyanobacteria and co-occurring algal groups, in particular dinoflagellates and co-occurring diatoms.[160] To examine the system performance of LDA and its capability to identify the phytoplankton samples at order level, a leave-one-out cross validation was conducted. The results are shown in Table 5-2. The algorithm works optimal, especially for cyanobacteria, as their spectral characteristics are significantly different to other algal groups. The accuracy for each algal group lies between 81.5%–100.0%. Specificity and false-positive rate also suggest a good performance of the algorithm, especially for the reliable identification of dinoflagellates and cyanobacteria with a low error rate (9.3% and 0.0%, respectively). The overall prediction accuracy of the ALPACA, when classifying into the groups of 'cyanobacteria', 'dinoflagellates' and 'other algae', is 97.3% (see the excel file in the supplemental material).

Table 5-2 System performance and figures of merit for each algal group when using eight LEDs within the ALPACA.[a]

phylum	accuracy	selectivity	sensitivity or TPR	specificity or TNR	FPR	FNR	error rate
cyanobacteria (1)[a]	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
dinoflagellates (11)[a]	90.7%	87.5%	63.6%	97.7%	2.3%	37.4%	9.3%
diatoms (11)[a]	81.5%	56.6%	46.5%	90.7%	9.3%	54.5%	18.5%
haptophytes (8)[a]	94.4%	100.0%	62.5%	100.0%	0.0%	37.5%	6.6%
chlorophytes (8)[a]	88.9%	62.5%	62.5%	93.5%	7.5%	37.5%	11.1%
ochrophytes (6)[a]	87.0%	33.3%	17.7%	96.8%	4.2%	83.3%	13.0%
cryptophytes (6)[a]	92.6%	100.0%	42.9%	100.0%	0.0%	57.1%	7.4%
euglenophytes (2)[a]	94.4%	0.0%	0.0%	98.1%	1.9%	100.0%	6.6%

[a] number of samples

The system performance of the algorithm was further tested against a full spectra analysis and a reduced LED selection using only four LEDs, which have the strongest impact on the discriminant vector according to Table S5-14. Detailed information about the assigned group membership for all evaluation variants and interferences between different algal groups are shown in the excel files of the supplemental material in Table S5-17. The overall prediction accuracy of both approaches – the full spectra analysis (92.0%) and the analysis with only four LEDs (98.1%) – is similar compared to the analysis with the entire LED selection.

5.5.2. Alterations in the Pigmentation during Growth and Under Different Light Irradiances

Pigment ratios are known to change for a variety of reasons, among others, light intensity, growth phase or nutrient availability.[50,51,161,162] These alterations have an important consequence for the reliability of pigment-based methods and might limit the scope of the multivariate discriminant analysis whose training database consists of marine phytoplankton samples cultivated under $100 \, \mu$mol / (s m²) and analyzed in the exponential phase. To examine the performance of the discriminant analysis under different growing conditions, changes in the relative pigmentation and its influence on the assigned group membership were evaluated (a) at different growth phases and (b) under different scalar irradiances. Alterations during growth are given in the supplemental material (Figure S5-11 to Figure S5-13).

Table 5-3 Calculated probability of group membership for the phytoplankton samples examined along the growth cycle and under different light conditions. [a]

species	Scalar irradiance (μmol / (s m2))	lag-phase real phylum	lag-phase identified phylum	exponential phase real phylum	exponential phase identified phylum	stationary phase real phylum	stationary phase identified phylum
E. huxleyi (haptophytes)	20	31.09	haptophytes*	23.65	dinoflagellates (26.27)	26.38	haptophytes
	100	19.64	chlorophytes (23.59)	27.37	haptophytes*	29.35	haptophytes
	200	19.0	chlorophytes (24.89)	27.61	haptophytes*	27.48	haptophytes
T. weissflogii (diatoms)	20	11.56	dinoflagellates (47.42)	22.8	diatoms	26.62	diatoms*
	100	23.13	diatoms	32.91	diatoms*	22.14	cryptophytes (30.51)
	200	20.08	ochrophytes (22.29)	13.77	ochrophytes (24.82)	17.01	ochrophytes (20.96)
Synechococcus sp. (cyano-bacteria)	20	98.48	cyanobacteria*	99.98	cyanobacteria*	99.98	cyanobacteria*
	100	99.62	cyanobacteria*	99.98	cyanobacteria*	99.96	cyanobacteria*
	200	99.76	cyanobacteria*	99.96	cyanobacteria*	100.00	cyanobacteria*

[a] sample was classified correctly and statistically secure

As shown in Table 5-3 all three phytoplankton species are classified correctly with statistical significance in the exponential phase, when cultivated under $100 \, \mu$mol / (s m²). For the cyanobacteria, the classification succeeds under every examined condition. As expected, for species with a pigmentation strongly depending on the light quality, intensity or growth phase, further training data would significantly improve the identification. Thus, the training data has to match

the intended application regarding the expected light conditions and biomass. In the future, this information may be provided by sensors, ideally as additional input parameters for the training database.

5.5.3. Mixed Sample Analysis

In the marine environment, algal blooms consist of a variety of different algal groups and species, which makes it indispensable to test the separation capability of the discriminant analysis in mixed samples.[163] As a first step towards real-world applications, we were interested at which cell density or cell ratio, the discriminant analysis is able to recognize the potential harmful cyanobacteria in mixed assemblages. LDA was applied to eleven mixtures of the prokaryotic cyanobacteria *Synechococcus* sp. and the eukaryotic diatom *T. weissflogii* to predict the probability of group membership based on the Mahalanobis distance. The predicted probability of group membership for each mixing ratio is given in Table 5-4.

Table 5-4 Cell density and mixing ratios for the diatom *T. weissflogii* and the cyanobacteria *Synechococcus* sp. and resulting probabilities of group membership after applying the linear discriminant analysis.

sample	cell density (cells per milliliter)		relative proportion	cyanobacteria : diatom ratio (cell density)	probability of group membership (%)	
	diatom	cyanobacteria			diatom	cyanobacteria
1	75,300	0	100 : 0	0	20.2	0.0
2	75,300	277,000	100 : 20	4	21.4	0.0
3	37,600	277,000	100 : 40	7	20.7	0.0
4	9,400	139,000	100 : 80	15	17.7	3.7
5	5,600	139,000	74 : 100	25	21.0	9.4
6	9,400	277,000	63 : 100	30	20.1	3.7
7	3,800	277,000	25 : 100	73	17.0	18.7
8	3,800	555,000	13 : 100	146	13.2	23.5
9	3,800	832,000	8 : 100	219	7.5	61.3
10	3,800	1387,000	5 : 100	365	7.1	69.7
11	0	1387,000	0 : 100	--	0.1	99.6

The determined probability of group membership refers to all algal groups in the training database, which is why the probability of the pure diatom sample cannot reach 100% due to the spectral overlap between diatoms and other algal groups. As shown in Table 5-4, the higher the relative proportion of cyanobacteria in the sample, the higher the predicted probability of cyanobacteria membership. When the diatom *T. weissflogii* dominates the sample composition in terms of total biomass and cell volume, cyanobacteria cannot be recognized by LDA. Starting from a mixing ratio of 100:80 (sample 4), LDA identifies cyanobacteria in the sample, although the predicted probability is low with 3.7%. When the mixing ratio is above 8:100, cyanobacteria dominate the

sample composition and the calculated probability for a cyanobacteria membership increases to 61.3%. Although LDA is sensitive for cyanobacteria, indicating the occurrence of the occurrence of cyanobacteria in mixed samples starting from a mixing ratio of ~15:1 (cell density). However, the separation capability of the discriminant analysis is highly affected by the absolute biovolume within the measurement channel. As the biovolume of the eukaryotic diatom is more than 20 times lager compared to the prokaryotic cyanobacteria, the cell density of cyanobacteria need to be much higher in order to identity them clearly in mixed samples.

Acknowledgments

5.6. Supporting Information

5.6.1. Algal Characterization

A total of 53 phytoplankton strains were obtained from the Culture Collection of the Plant biology and Ecology Department from the University of the Basque Country (UPV/EHU) in Spain. The strains belong to 8 different phyla and 27 orders. Each strain was cultured in conical flasks with f/2 growth medium at a salinity of 33 and with addition of selenium and silicate as appropriate. The cultivating temperature was 17 °C and the irradiances with a cool white fluorescent tube were 20, 100 or 200 μmol / (s·m²). The light:dark cycle during cultivation was 10:14 h.

Table S5-1 Marine phytoplankton species studied and scalar irradiance for cultivation in batch culture. [a] number of species examined in round parentheses.

phylum [a]	class	order	species and strain	scalar irradiance (μmol / (s·m²))	abbrev.
Cyanobacteria (1)	Cyanophyceae	Synechococcales	*Synechococcus* sp.	20, 100 & 200	Syn
Dinophyta (dinoflagellates) (11)	Dinophyceae	Gonyaulacales	*Alexandrium cf. hiranoi (Dn5EHU)*	100	Ale1
			Alexandrium mediterraneum	100	Ale2
			Alexandrium minutum (Dn5EHU)	100	Ale3
		Gymnodiniales	*Amphidinium* sp.	100	Amp
			Karlodinium veneficum (Dn43EHU)	100	Kar
			Lepidodinium chlorophorum (Dn16EHU)	100	Lep
		Peridiniales	*Heterocapsa cf. pygmaea (Dn7EHU)*	100	Het
			Kryptoperidinium foliaceum	100	Kry
			Cf. Scrippsiella (Dn1EHU)	100	Scr
		Prorocentrales	*Prorocentrum lima*	100	Pro1
			Prorocentrum triestrium	100	Pro2
Bacillariophyta (diatoms) (11)	Bacillariophyceae	Bacillariales	*Cylindrotheca closterium*	100	Cyl
	Mediophyceae	Stephanodiscales	*Cyclotella atomus (Bc9EHU)*	100	Cyc1

			Cyclotella cryptica (Bc1EHU)	100	Cyc2
			Cyclotella meneghiniana (Bc11EHU)	100	Cyc3
		Chaetocerotales	*Chaetoceros curvisetus*	100	Cha1
			Chaetoceros sp.	100	Cha2
		Cymatosirales	*Minutocellus cf. polymorphus*	100	Min
		Thalassiosirales	*Skeletonema costatum (Bc24EHU)*	100	Ske1
			Skeletonema dohrnii (Bc20EHU)	100	Ske2
Haptophyta (8)			*Thalassiosira weissflogii (Bc2EHU)*	20, 100 & 200	Tha
			Thalassiosira pseudonana (Bc4EHU)	100	Tha1
	Coccolithophyceae	Coccolithales	*Pleurochrysis roscoffensis (Hp19ehu)*	100	Ple
		Isochrysidales	*Emiliania huxleyi*	20, 100 & 200	Emi
		Prymnesiales	*Haptolina hirta (Hp54EHU)*	100	Hap
			Imantonia rotunda (Hp2EHU)	100	Ima
			Isochrysis galbana (Hp1EHU)	100	Iso
			Prymnesium palpebrale (Hp11ehu)	100	Pry
			Chrysochromulina throndsenii (Hp4ehu)	100	Chr
Chlorophyta (8)	Chlorophyceae	Chlamydomonadales	*Chlamydomonas cf. coccoides (CGA20)*	100	Chl
	Chlorodendrophyceae	Chlorodendrales	*Tetraselmis sp.*	100	Tet
	Ulvophyceae	Oltmannsiellopsidales	*Oltmannsiellopsis viridis (CGA18)*	100	Olt
	Mamiellophyceae	Mamiellales	*Mamiella gilva*	100	Mam
			Micromonas pusilla	100	Mic
			Ostreococcus sp.	100	Ost
	Nephrophyceae	Nephroselmidales	*Nephroselmis pyriformis (CGA7)*	100	Nep

	Pyramimonadophyceae	Pyramimonadales	*Pyramimonas orientalis (CGA14)*	100	Pyr
Ochrophyta (6)	Raphidophyceae	Chattonellales	*Heterosigma akashiwo (Rp1EHU)*	100	Het
	Chrysophyceae	Chromulinales	*Ochromonas sp. (Cs1EHU, Cs2EHU)**	100	Och1 / Och2
			Ollicola vangoorii (Cs3EHU)	100	Oll
	Eustigmatophyceae	Eustigmatales	*Nannochloropsis* sp. *(Es1EHU)*	100	Nan
	Pelagophyceae	Sarcinochrysidales	*Ankylochrysis lutea (Pl2EHU)*	100	Ank
Cryptophyta (6)	Cryptophyceae	Cryptomonadales	*Chroomonas collegionis (Cr36EHU)*	100	Chr1
			Chroomonas sp. (Cr83EHU)	100	Chr2
			*Hemiselmis cf. rufescens (Cr21EHU, Cr31EHU) **	100	Hem1 / Hem2
			Hemiselmis cf. tepida (Cr38EHU)	100	Hem3
		Pyrenomonadales	*Teleaulax gracilis (Cr6EHU)*	100	Tel
			Urgorri complanatus (Cr1EHU)	100	Urg
Euglenophyta (2)	Euglenophyceae	Eutreptiales	*Eutreptiella gymnastica*	100	Eut1
			Eutreptia cf. viridis	100	Eut2

[a] Two different strains

5.6.2. HPLC Analysis of Major Pigments

Feature extraction was the initial step within the discriminant analysis as it is to any pattern classification technique. It aims at reducing dimensionality allowing a reliable differentiation of all algal classes with a minimum number of spectral properties. In order to bind spectral characteristics, found in the fluorescence excitation spectra, to the pigment-based taxonomic principle, we investigated major pigments of each algal class by HPLC analysis. Relevant marker pigments were linked to spectral differences of algae, facilitating the extraction of spectral key features.

However, according to Bidigare et al.[23] the overall absorption by algae is not just an accumulation of absorption coefficients of individual pigments. A direct link between pigments examined by HPLC and excitation spectra is not possible, due to specific pigment packaging effects present in the cell. These packaging effects and pigment-protein complexes cause alterations in the individual absorption properties of pigments, leading to a 'shifting' in the overall absorption spectra. Furthermore, variations in the pigment concentrations can arise from different light conditions, nutrient availability or during growth.[10,50,72]

Besides Bidigare, also Bricaud[72] deals with theoretical reconstruction and simulations to match in-vivo absorption of phytoplankton. They assume that the overall phytoplankton absorption a_{overall} comprises a linear combination of the in-vivo weight-specific absorption coefficient of the pigments and their intrinsic concentration: $a_{\mathrm{overall}}(\lambda) = \sum a_i(\lambda) \cdot c_i$. Based on their studies, they established a database with theoretical absorption coefficients ($400 - 700\,\mathrm{nm}$) for major (non)-photosynthetically active pigments present in algal cells.[23,72]

After combining the recorded fluorescence excitation spectrum with theoretical absorption coefficients of major pigments, we were able to link spectral differences between different algal classes with evaluated marker pigments. Besides allocation of spectral characteristics to major pigments for each algal class, we validated the relative fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectra recorded by the spectrofluorometer.

5.6.2.1. Cyanobacteria
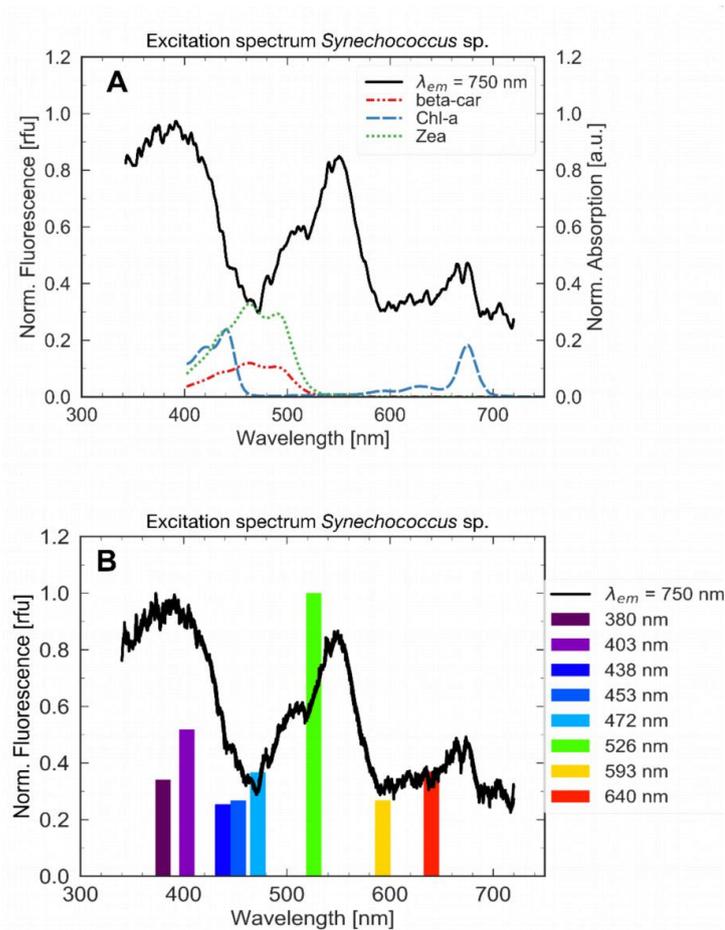
Figure S5-1: (A) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the cyanobacteria *Synechococcus* sp.. **(B)** Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-2 Pigment : chlorophyll-*a* molar ratios of the cyanobacteria *Synechococcus* sp. studied after growing under $100\,\mu\text{mol}\,/\,(\text{s}\,\text{m}^2)$ and analyzed in the exponential phase. Phycobilines cannot be analyzed by the conducted HPLC analysis.

	Zea	$\beta\beta$-Car	Chl-*a*
Synechococcus sp.	1.39	0.49	1.00

5.6.2.2. Dinoflagellates
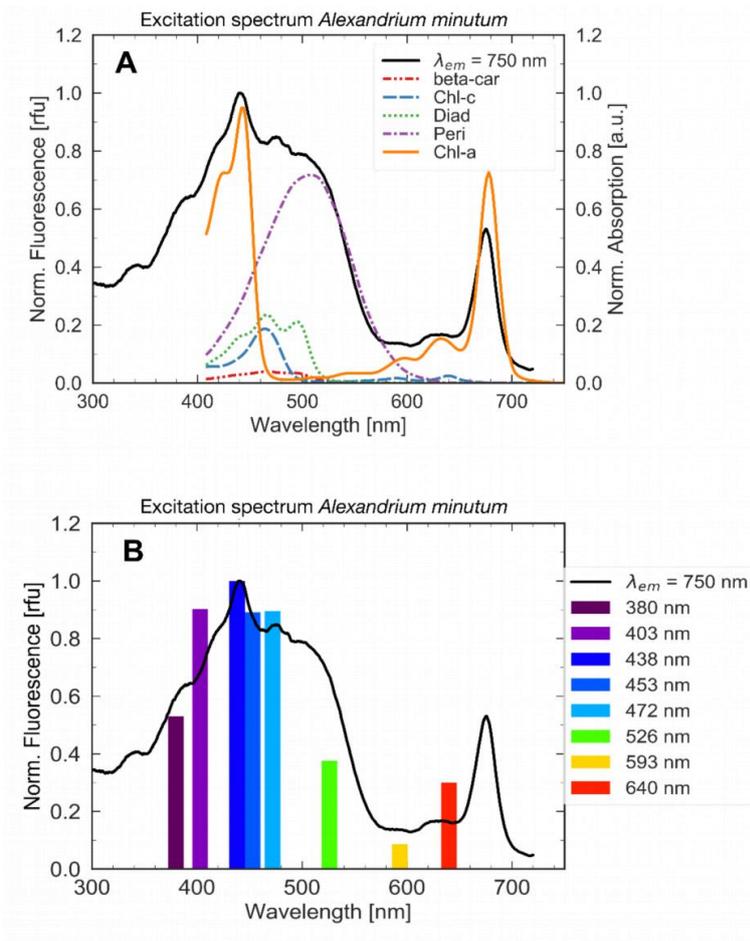
Figure S5-2: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the dinoflagellate *Alexandrium minutum* (A. *minutum*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-3 Pigment : chlorophyll-*a* molar ratios of the dinoflagellates studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Chl-c3	MgDVP	Chl-c2	Chl-c1	Per	19BF	Fuco	Cis-Neo	Viola	19HF	Diadchr	Diadino	Dino	Asta	Anth	Diato	Zea	Lut	GyrE	GyrE	Chl-b	Chl-a	βψ-Car	βε-Car	ββ-Car
Ale1	--	--	0.28	--	0.91	--	--	--	--	--	0.03	0.17	0.08	--	--	--	--	--	--	--	--	1.00	--	--	0.03
Ale2	--	--	0.19	--	0.84	--	--	--	--	--	0.02	0.20	0.06	--	--	0.14	--	--	--	--	--	1.00	--	--	0.05
Ale3	--	--	0.19	--	0.72	--	--	--	--	--	0.04	0.24	0.06	--	--	--	--	--	--	--	--	1.00	--	--	0.04
Amp	--	0.02	0.54	--	0.89	--	--	--	--	--	0.01	0.34	0.08	--	--	0.04	--	--	--	--	--	1.00	--	--	0.04
Kar	0.07	--	0.18	--	--	0.10	0.40	--	--	0.26	--	0.07	--	--	--	0.07	--	--	0.13	0.01	--	1.00	--	--	0.02
Lep	--	--	--	--	--	--	--	0.11	0.13	--	--	--	--	0.0	0.0	--	0.0	0.11	--	--	0.5	1.0	--	0.0	0.0
Het1	--	0.01	0.39	--	1.00	--	--	--	--	--	0.01	0.26	0.08	--	--	0.01	--	--	--	--	--	1.00	--	--	0.04
Kry	--	--	0.03	0.10	--	--	0.75	--	--	--	--	0.12	--	--	--	--	--	--	--	--	--	1.00	0.02	--	0.06
Scr	--	--	0.36	1.04	--	--	--	--	--	--	0.02	0.24	0.04	--	--	0.01	--	--	--	--	--	1.00	--	--	0.04
Pro1	--	--	0.13	--	0.94	--	--	--	--	--	0.38	0.13	0.02	--	--	0.02	--	--	--	--	--	1.00	--	--	0.03
Pro2	--	--	0.29	--	0.82	--	--	--	--	0.25	--	--	--	--	--	--	--	--	--	--	--	1.00	--	--	0.04

5.6.2.3. Diatoms

Figure S5-3: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the diatom *Thalassiosira weissflogii* (T. *weissflogii*) or *Conticribra weissflogii*, (C. *weissflogii*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-4 Pigment : chlorophyll-*a* molar ratios of the diatoms studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Chlyde-*a*	MgDVP	Chl-*c2*	Chl-*c1*	Fuco	Diadino	Diato	Chl-*a* allomer	Chl-*a*	Chl-*a*'	$\beta\epsilon$-Car	$\beta\beta$-Car
Cyl	--	--	0.12	0.04	0.60	0.04	0.53	--	1.00	--	--	0.04
Cyc1	--	--	0.04	0.05	0.44	0.06	0.53	--	1.00	--	--	0.05
Cyc2	--	--	0.05	0.06	0.48	0.86	0.53	--	1.00	--	--	0.05
Cyc3	--	--	0.05	0.03	0.38	0.12	0.53	--	1.00	--	--	0.07
Cha1	1.10	--	4.23	0.91	2.58	0.07	0.53	--	1.00	--	--	0.18
Cha2	0.10	--	0.15	0.16	0.68	0.02	0.53	--	1.00	--	--	0.05
Min	--	--	0.14	0.07	0.70	0.07	0.53	--	1.00	--	--	0.05
Con	--	0.01	0.04	0.03	0.38	0.18	0.01	0.01	1.00	0.01	0.05	4.90
Ske1	0.05	--	0.12	0.07	0.66	0.01	0.53	--	1.00	--	--	0.04
Ske2	0.35	--	0.39	0.16	0.70	0.03	0.53	--	1.00	--	--	0.07
Tha	--	--	0.05	0.07	0.48	0.05	0.53	--	1.00	--	--	0.04

5.6.2.4. Haptophytes
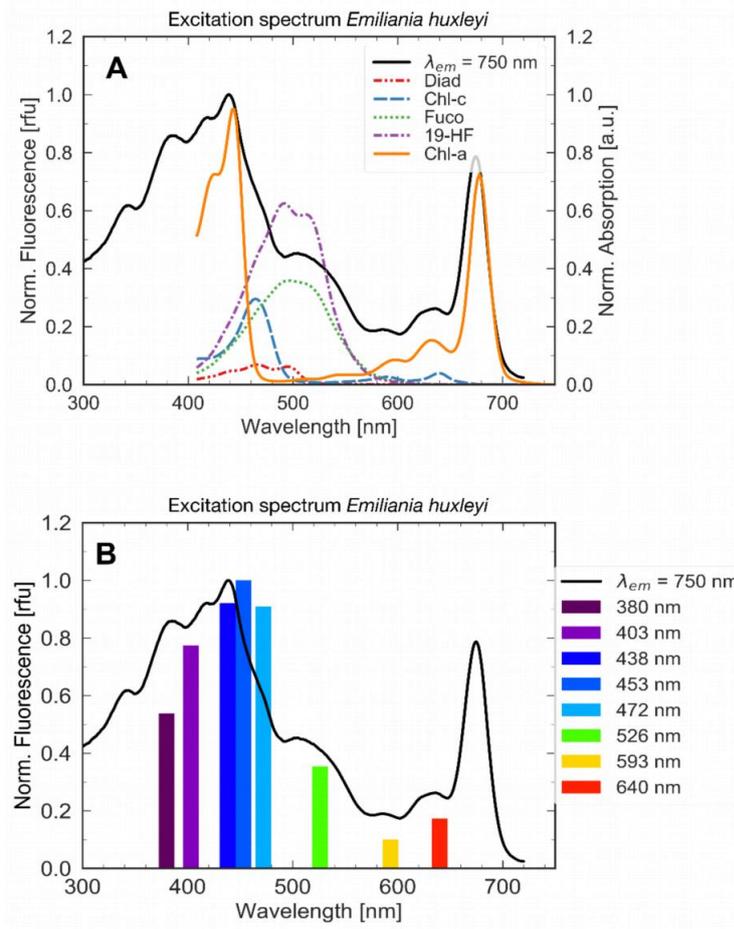
Figure S5-4: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the haptophyte *Emiliania huxleyi* (E. *huxleyi*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-5 Pigment : chlorophyll-*a* molar ratios of the haptophytes studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Chl-c2 pg	Chl-c3	MgDVP	Chl-c2	Chl-c1	19BF	Fuco	4KHF	19HF	Diadino	Diato	Gyro	Np Chl-c2	Np Chl-c2 18:14 14:00	Chl-a	βε-Car	ββ-Car
Ple	--	--	--	0.02	0.04	--	0.27	--	--	0.17	0.02	--	--	--	1.00	--	0.05
Emi	--	0.16	0.01	0.30	--	--	0.36	0.10	0.63	0.07	0.07	0.01	0.13	--	1.00	--	0.04
Hap	--	0.18	0.04	0.19	--	--	0.25	0.31	0.47	0.21	0.03	0.02	0.01	0.16	1.00	0.01	0.03
Ima	--	0.25	0.02	0.20	--	0.17	0.18	0.09	0.89	0.19	0.05	0.02	0.01	0.11	1.00	0.01	0.03
Iso	--	--	0.01	0.31	0.13	--	0.71	--	--	0.23	0.02	0.01	--	0.07	1.00	--	0.05
Pry	--	0.19	0.01	0.21	--	--	1.03	0.07	0.03	0.14	0.03	0.01	--	0.09	1.00	0.01	0.02
Pav	0.06	--	0.01	0.04	0.08	--	0.56	--	--	0.21	0.02	--	--	--	1.00	--	0.07

5.6.2.5. Chlorophytes

Chlorophytines
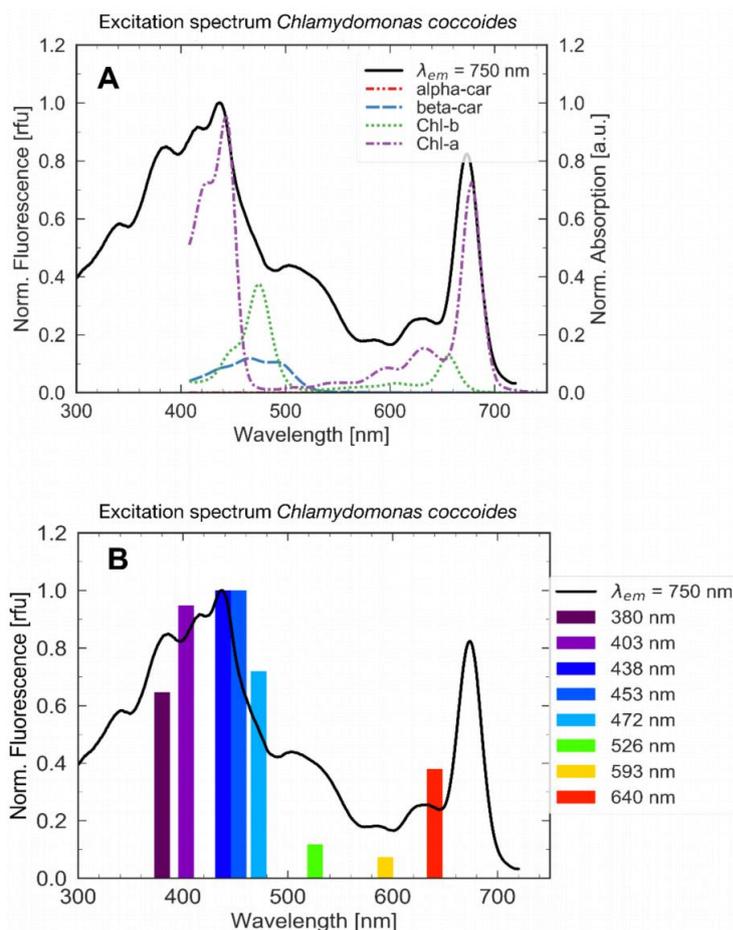
Figure S5-5: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the chlorophytine *Chlamydomonas coccoides* (C. *coccoides*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-6 Pigment : chlorophyll-*a* molar ratios of the chlorophytines studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Cis-Neo	Viola	Anth	Lut	Chl-*b*	Chl-*a*	$\beta\psi$-Car	$\beta\beta$-Car
Chl	0.06	0.06	--	0.21	0.39	1.00	0.03	0.12
Tet	0.11	0.09	--	0.12	0.67	1.00	0.02	0.12
Olt	0.11	0.20	0.01	0.19	0.66	1.00	--	0.13

Multiparametric Sensing – from Development to Application in Marine Science
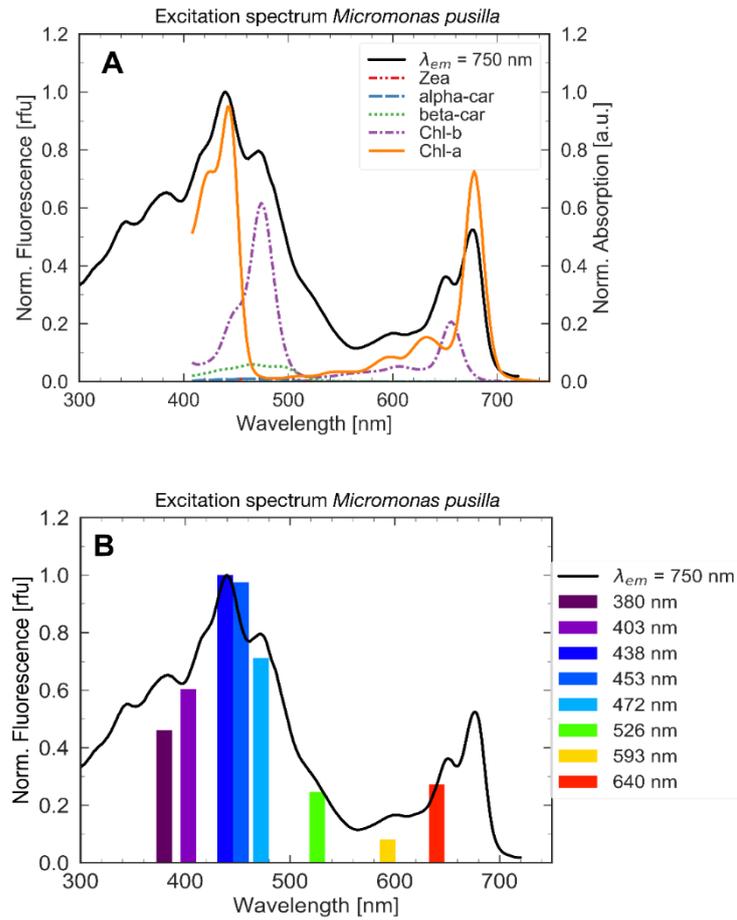
Prasinophytina

Figure S5-6: (A) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the prasinophytine *Micromonas pusilla* (M. *pusilla*). **(B)** Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-7 Pigment : chlorophyll-a molar ratios of the prasinophytes studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	MgDVP	Uri	Cis-Neo	Pra	Mic	Viola	Micro	Anth	Zea	Lut	Dihi	Siph	Chl-*b*	Chl-*a*	$\beta\psi$-Car	$\beta\varepsilon$-Car	$\beta\beta$-Car
Mam	0.06	0.08	0.11	0.22	0.04	0.17	0.05	-	0.04	0.02	0.04	-	0.72	1.00	-	0.03	0.17
Mic	0.09	0.09	0.11	0.17	0.07	0.17	0.07	-	0.01	-	0.02	-	0.64	1.00	-	0.01	0.06
Ost	0.07	0.06	0.10	0.18	-	0.17	0.04	0.01	0.02	0.07	-	-	0.50	1.00	-	-	0.06
Nep	0.07	-	0.11	-	-	0.10	-	0.04	0.02	0.02	-	0.02	0.68	1.00	-	0.06	0.06
Pry	0.02	-	0.13	-	-	0.21	-	0.02	0.02	0.07	-	0.10	0.68	1.00	0.03	0.01	0.19

Multiparametric Sensing – from Development to Application in Marine

5.6.2.6. Cryptophytes

Figure S5-7: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the cryptophyte *Chroomonas collegionis* (*C. collegionis*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-8 Pigment : chlorophyll-*a* molar ratios of the cryptophytes studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase. Phycobilines cannot be analyzed by the conducted HPLC analysis.

	Chlyde-*a*	Chl-*c2*	Allo	Croco	Chl-*a*	$\beta\Psi$-Car	$\beta\varepsilon$-Car	$\beta\beta$-Car
Chr1	--	0.05	0.62	0.04	1.00	0.14	0.09	0.12
Chr2	0.02	0.08	0.40	0.03	1.00	--	0.08	--
Hem1	--	0.18	0.55	0.06	1.00	--	0.06	--
Hem2	0.04	0.14	0.68	0.05	1.00	--	0.08	--
Hem3	--	0.15	0.42	0.04	1.00	--	0.05	--
Tel	--	0.14	0.46	0.04	1.00	--	0.07	--
Urg	0.02	0.15	0.56	0.07	1.00	--	0.11	--

5.6.2.7. Ochrophytes
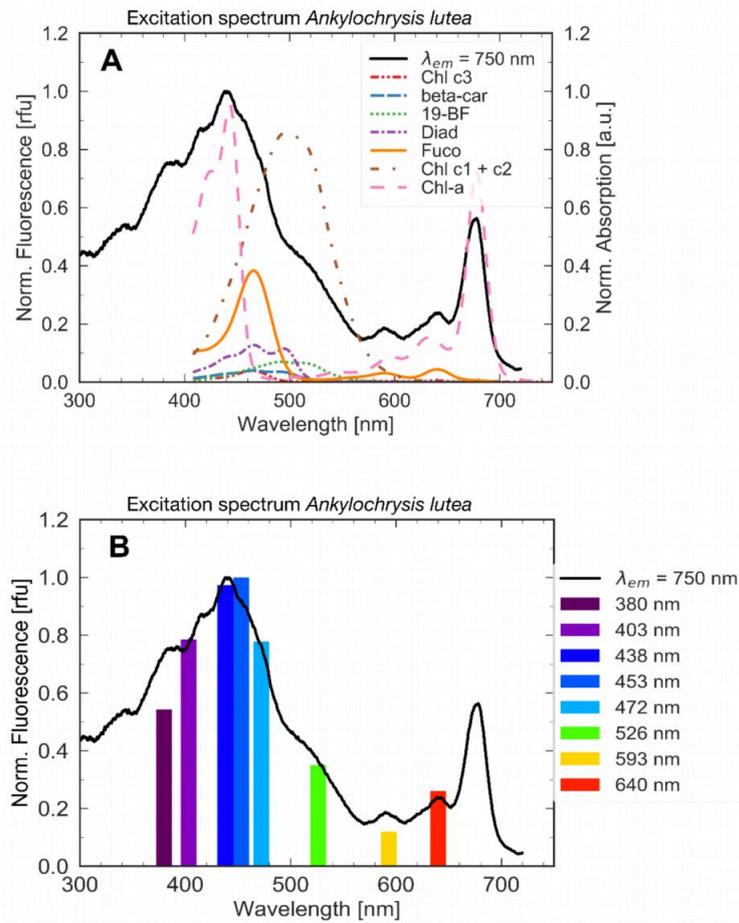
Figure S5-8: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the ochrophyte *Ankylochrysis lutea* (A. *lutea*). (**B**) Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-9 Pigment : chlorophyll-*a* molar ratios of the ochrophytes studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Chl-*c3*	MgDVP	Chl-*c2*	Chl-*c1*	Vauch	19BF	Fuco	Viola	Diadino	Anth	Diato	Zea	Vauch-ester	Chl-*a*	*ββ*-Car
Het	--	0.01	0.13	0.01	--	--	0.52	0.19	--	0.01	--	0.01	--	1.00	0.06
Och1	--	--	--	0.06	--	--	0.25	0.06		--	--	--	--	1.00	0.10
Och2	--	--	--	--	--	--	0.31	--	--	--	--	--	--	1.00	--
Oll	0.04	0.01	0.39		--	0.07	0.86	0.01	0.13	--	0.03		--	1.00	0.04
Nan	--	--	--	--	0.02	--	--	0.39	--	--	--	0.25	0.03	1.00	0.06
Ank	--	0.01	0.13	0.01	--	--	0.52	0.19	--	0.01	--	0.01	--	1.00	0.06

5.6.2.8. Euglenophytes
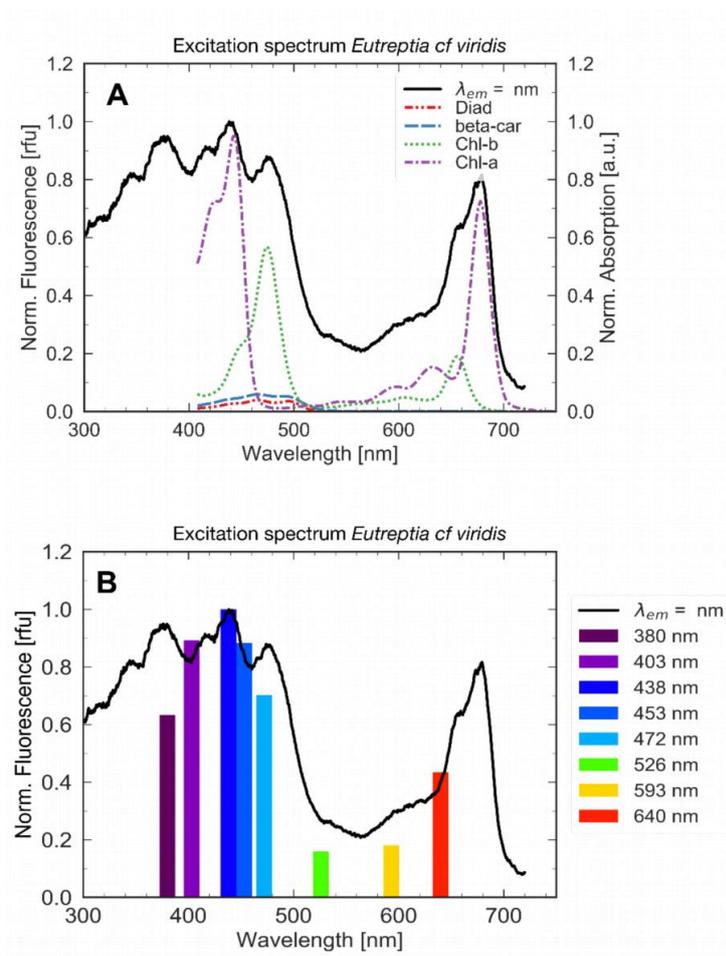
Figure S5-9: (**A**) Comparison of the fluorescence excitation spectrum with theoretical absorption coefficients of major pigments to allocate spectral characteristics of the euglenophyte *Eutreptia viridis* (E. *viridis*). (**B**). Validation of the fluorescence pattern recorded by the algae detection module with the fluorescence excitation spectrum acquired with a fluorescence spectrofluorometer.

Table S5-10 Pigment : chlorophyll-*a* molar ratios of euglenophytes studied after growing under 100 μmol / (s m²) and analyzed in the exponential phase.

	Cis-Neo	Diadino	Diato	Siph	Eutrept	Chl-*b*	Chl-*a*	$\beta\beta$-Car
Eut1	0.10	0.04	0.04	0.08	0.11	0.59	1.00	0.06
Eut2	0.10	0.51	0.06	0.05	0.18	0.60	1.00	0.07

5.6.3. Operational Features of the LDA

Statistical information of the training database is given below. First, the input matrix for LDA is listed in Table S5-11. Herein, the rows list the phytoplankton strains which are used as training objects. These strains are excited at eight different wavelengths. The entries of each column, known as measurement features, correspond with the auto-scaled fluorescence intensity, that was measured by the ALPACA upon excitation. These input data are grouped at order level and analyzed by LDA which computes the scores of each strain. Then, strains of the same phylum are combined, subsequently called group centroids. For clarification, Figure S5-10 provides an overview of the process steps during data evaluation. The rows in Table S5-12 summarize the standard deviation of each group centroid in each projection direction. In Table S5-13, the eigenvalues used for discrimination and their contribution to the explained variance of the original data are shown for different taxonomic levels.
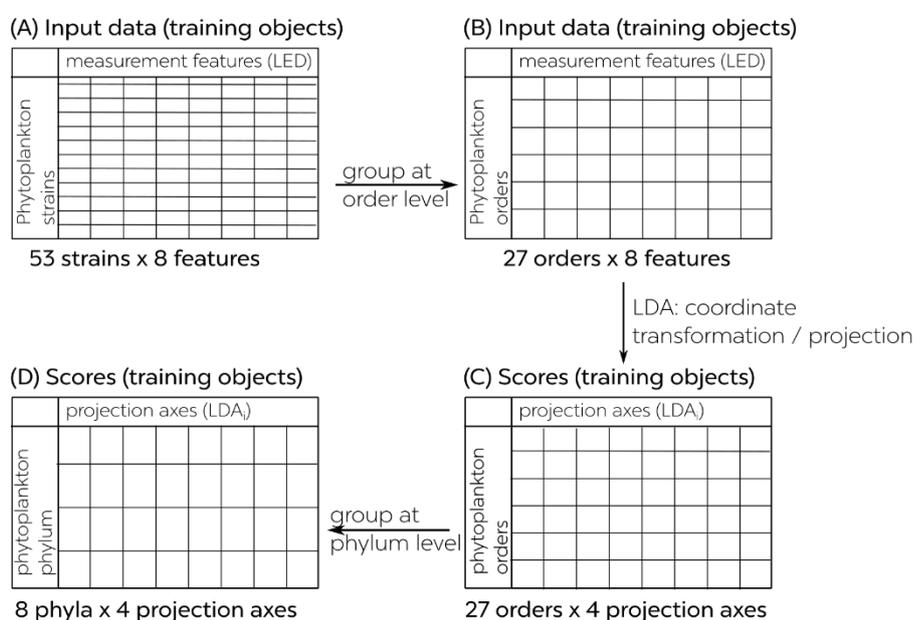
Figure S5-10: Working principle of LDA and description on how the input matrix changes along the process steps. (**A**) Initial input matrix, where the phytoplankton strains used as training objects are listed in rows. Along the columns, the measurement features (excitation wavelengths) are listed. (**B**) In the first step, training objects are groups at order level as the separation works best in this case. (**C**) Then, LDA is applied to compute the scores of the training objects. These scores are finally grouped again at phylum level (**D**).

Table S5-11 Input matrix for LDA. The columns indicate the excitation wavelengths, also known as measurement features, and in the row the phytoplankton strains (training objects) are listed. The entries are the cumulative light intensity upon excitation at corresponding LED recorded by the ALPACA. The light intensity is auto-scaled.

phytoplankton abbrev.	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
Syn	-0.239	0.327	-0.417	-0.434	-0.104	2.307	-0.915	-0.526
Ale1	-0.87	-1.014	0.412	1.731	1.044	-0.758	0.069	-0.614
Ale2	-0.537	-0.27	2.162	0.247	0.174	0.148	-0.969	-0.956
Ale3	-0.7035	-0.642	1.287	0.989	0.609	-0.305	-0.45	-0.785
Amp	-0.331	-0.65	1.247	1.522	0.605	-0.349	-1.121	-0.923
Kar	-0.063	0.831	0.989	0.908	0.701	-0.763	-1.444	-1.159
Lep	0.008	0.959	0.873	0.752	0.655	-1.325	-1.563	-0.359
Het	-0.156	0.698	0.855	1.168	0.645	-0.51	-1.537	-1.163
Kry	-0.361	1.007	1.857	0.151	-0.678	-0.213	-1.287	-0.476
Scr	-0.411	1.062	0.981	0.825	0.577	-0.473	-1.616	-0.946
Pro1	-0.14	0.556	1.121	0.971	0.73	-0.583	-1.525	-1.129
Cyl	-0.041	1.427	0.805	0.649	0.217	-0.795	-1.702	-0.56
Cyc1	0.068	1.298	0.991	0.892	-0.167	-0.618	-1.443	-1.021
Cyc2	-0.062	1.529	1.14	0.453	-0.316	-0.704	-1.509	-0.532
Cyc3	-0.053	1.045	1.288	0.86	-0.037	-0.762	-1.476	-0.865
Cha1	-0.096	0.971	1.11	1.028	0.202	-0.769	-1.482	-0.964
Cha1	-0.113	1.208	1.067	0.738	0.233	-0.732	-1.632	-0.77
Min	-0.151	0.783	1.114	1.143	0.235	-0.584	-1.582	-0.959
Ske1	-0.21	0.936	1.202	0.919	0.321	-0.666	-1.487	-1.015
Ske2	-0.176	1.064	1.168	0.76	0.355	-0.658	-1.557	-0.955
Tha	0.044	1.5	0.892	0.785	-0.369	-0.615	-1.478	-0.755
Tha1	0.007	0.973	1.249	0.866	-0.007	-0.603	-1.625	-0.86
Ple	-0.074	1.199	0.847	0.897	-0.129	-1.302	-1.522	0.084
Emi	-0.143	0.638	0.901	1.127	0.736	0.627	-1.505	-1.129
Hap	-0.066	0.454	0.875	1.17	0.857	-0.635	-1.419	-1.232
Ima	-0.147	0.403	0.932	1.225	0.768	-0.497	-1.394	-1.291
Iso	-0.155	0.404	1.078	1.278	0.569	-0.601	-1.389	-1.185
Pry	-0.057	0.851	0.798	1.016	0.701	-0.571	-1.509	-1.229
Chr	-0.108	0.695	0.786	1.133	0.805	-0.667	-1.411	-1.232
Pav	-0.089	0.799	1.072	1.155	0.265	-0.592	-1.459	-1.15
Chl	0.088	0.878	1.025	1.038	0.279	-1.292	-1.405	-0.612
Tet	-0.13	0.927	1.026	1.036	0.205	-1.244	-1.523	-0.296
Olt	0.089	1.171	0.81	0.778	0.324	-1.376	-1.51	-0.286
Mam	-0.087	0.338	1.153	1.193	0.659	-0.822	-1.382	-1.052
Mic	-0.244	0.178	1.342	1.267	0.494	-0.874	-1.363	-0.799
Ost	-0.124	0.257	1.303	1.253	0.461	-0.871	-1.344	-0.935

Multiparametric Sensing – from Development to Application in Marine Science

Nep	-0.224	0.42	1.166	1.195	0.583	-0.829	-1.535	-0.776
Pyr	-0.043	0.567	1.143	1.117	0.51	-1.026	-1.479	-0.789
Het	0.007	1.221	1.065	0.767	0.08	-0.588	-1.573	-0.98
Och1	-0.152	0.891	1.231	1.328	-0.545	-0.684	-1.035	-1.035
Och2	-0.081	0.856	1.207	0.992	0.239	-0.712	-1.505	-0.996
Oll	-0.065	0.574	1.239	1.233	0.2	-0.725	-1.355	-1.1
Nan	0.188	0.848	1.468	0.877	-0.283	-1.028	-1.183	-0.886
Ank	-0.127	0.521	1.083	1.198	0.564	-0.676	-1.429	-1.135
Chr1	-0.272	0.723	0.721	0.264	-0.801	-1.463	-0.745	1.574
Chr2	-0.434	0.522	0.428	0.053	-0.983	-1.211	-0.328	1.952
Hem1	-0.183	0.922	1.297	1.109	-0.116	-1.164	-1.154	-0.711
Hem2	0.114	0.909	0.992	0.882	0.13	-0.068	-1.48	-1.48
Hem3	-0.21	0.922	1.235	0.899	0.115	-0.271	-1.523	-1.168
Tel	0.145	0.983	0.827	0.908	0.447	-0.626	-1.736	-0.948
Urg	-0.066	1.028	1.186	0.905	0.083	-0.524	-1.408	-1.204
Eut1	0.103	0.906	1.212	0.818	0.189	-1.239	-1.485	-0.505
Eut2	0.07	0.878	1.189	0.848	0.283	-1.396	-1.328	-0.545
Syn	-0.239	0.327	-0.417	-0.434	-0.104	2.307	-0.915	-0.526
Ale1	-0.87	-1.014	0.412	1.731	1.044	-0.758	0.069	-0.614
Ale2	-0.537	-0.27	2.162	0.247	0.174	0.148	-0.969	-0.956

Table S5-12 Standard deviation of the group centroids of the training database analyzed by LDA.

	LDA1	LDA2	LDA3	LDA4
cyanobacteria	0.75	0.67	0.63	0.32
dinoflagellates	0.60	1.30	1.11	0.82
bacillariophytes (diatoms)	0.41	0.58	0.73	0.53
haptophytes	0.21	0.40	0.52	0.27
chlorophytes	0.25	0.39	0.25	0.26
ochrophytes	0.22	0.51	0.40	0.90
cryptophytes	1.92	1.21	0.70	1.46
euglenophytes	0.08	0.70	1.04	0.08

Table S5-13 Eigenvalues used for the discriminant analysis and their contribution to the explained variance. [a]

Taxonomic level [a]	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	sum of λ_{1-3}
phylum (8)	92.51%	3.70%	1.64%	1.06%	0.58%	0.37%	0.14%	97.85
class (18)	88.66%	6.19%	2.82%	1.56%	0.97%	0.60%	0.20%	97.67
order (27)	96.35%	3.80%	0.77%	0.08%	0.00%	0.00%	0.00%	99.92
family (42)	88.30%	7.30%	2.40%	1.44%	0.46%	0.10%	0.00%	98.00

[a] number of groups

When the first three eigenvalues $\lambda_{1\text{-}3}$ and their corresponding discriminant vectors w_{1-3} are selected, the explained variance of the original data is maximal at order level (99.92%). Thus, the separation is conducted at this level, however the group membership is returned at phylum level to enhance the figures of merit for the LDA.

According to equation 5.1 of the main part, the vector of discriminants w represents a weight linear combination of measurement features φ_i, here the fluorescence intensities upon excitation at distinct excitation wavelengths. The weighting factors α_i or scales are given in Table S5-14. The higher the absolute value, the more important the corresponding measurement feature for separation: For the first projection axis (LDA$_1$), the relative light intensity upon excitation at 526 nm and 453 nm has the strongest impact for the separation along the first projection axis, whereas the excitation wavelengths 380 nm or 640 nm are less relevant.

Table S5-14 Scaling of the discriminating variables for the projection axes (discriminant vectors).

	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
LDA 1	0.05	0.36	2.45	3.18	-1.30	-3.97	-2.55	0.35
LDA 2	-10.43	-0.18	-2.27	2.60	-0.17	0.96	-7.84	1.66
LDA 3	1.01	-2.01	-1.66	4.19	-4.22	1.38	-7.24	1.74
LDA 4	-7.11	2.12	0.46	1.42	-2.31	0.87	0.16	-0.23

Pearson's correlation matrix, given in Table S5-15, shows the correlation coefficients between sets of excitation wavelengths. Each of the measurement features is correlated with each of the other features. The correlation coefficients indicate, whether there is a dependency between two measurement features and, if they are dependent, whether the dependency is positive or negative.

Table S5-15 Correlation matrix of training objects indicating linear (in)dependency of spectral features. [a]

	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
380 nm	1.00							
403 nm	0.65 [a]	1.00						
438 nm	0.11	0.12	1.00					
453 nm	0.08	-0.18	0.48	1.00				
472 nm	-0.06	-0.42	0.04	0.59	1.00			
526 nm	-0.25	-0.26	-0.60	-0.63 [a]	-0.10	1.00		
593 nm	-0.69 [a]	-0.62 [a]	-0.37	-0.25	-0.27	0.25	1.00	
640 nm	-0.24	0.04	-0.27	-0.48	-0.60 [a]	-0.16	0.50	1.00

[a] correlation is significant

Results of the leave-one-out cross validation and independent figures of merits are shown in Table S5-16 Besides the full spectra analysis, a reduced analysis, when only four measurement features of the ALPACA are used for separation, are presented.

Table S5-16 (a) System performance and figures of merit for each algal group a full spectra analysis is conducted. [a] number of samples analyzed.

phylum [a]	accuracy	selectivity	sensitivity or TPR	specificity or TNR	FPR	FNR	error rate
cyanobacteria (1)	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
dinoflagellates (9)	90.0%	76.0%	67.7%	96.1%	4.9%	33.3%	10.0%
diatoms (11)	90.0%	87.5%	63.6%	97.4%	2.6%	37.4%	10.0%
haptophytes (8)	94.0%	100.0%	62.5%	100.0%	0.0%	37.5%	7.0%
chlorophytes (8)	98.0%	88.9%	100.0%	97.6%	2.4%	0.0%	2.0%
ochrophytes (5)	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
cryptophytes (6)	92.0%	76.0%	50.0%	97.7%	2.3%	50.0%	8.0%
euglenophytes (2)	98.0%	100.0%	50.0%	100.0%	0.0%	50.0%	2.0%

(b) System performance and figures of merit for each algal group when using only four LEDs within the ALPACA (380 nm, 438 nm, 453 nm and 526 nm).[a]

Phylum [a]	Accuracy	Selectivity	Sensitivity or TPR	Specificity or TNR	FPR	FNR	Error Rate
Cyanobacteria (1)	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
Dinoflagellates (9)	87.0%	83.3%	46.5%	97.7%	2.3%	54.5%	13.0%
Diatoms (11)	79.6%	--	0.0%	100.0%	0.0%	100.0%	20.4%
Haptophytes (8)	92.6%	83.3%	62.5%	97.8%	2.2%	37.5%	7.4%
Chlorophytes (8)	90.7%	100.0%	37.5%	100.0%	0.0%	62.5%	9.3%
Ochrophytes (5)	87.0%	40.0%	33.3%	93.8%	7.3%	67.7%	13.0%
Cryptophytes (6)	87.0%	--	0.0%	100.0%	0.0%	100.0%	13.0%
Euglenophytes (2)	97.3%	50.0%	100.0%	97.2%	3.8%	0.0%	3.7%

[a] number of samples analyzed

5.6.4. Relative Pigmentation along Growth Cycle and Under Different Light Conditions

Relative pigmentation within the photosystems of phytoplankton change for a variety of reasons, among other, during growth, reflecting nutritional statues, or due to different light conditions. These alterations might have an important effect on the reliability of pigment-based methods and the discriminant analysis of choice. To examine the performance of the discriminant analysis under abnormal conditions, changes in the relative pigmentation investigated by HPLC analysis and its effect on the predicted group membership were evaluated in three representatives of the major algal classes haptophytes (*Emiliania huxleyi*), diatoms (*Thalassiosira weissflgoii*) and cyanobacteria (*Synechococcus* sp.). The allocated group membership is taken as statistically secure, if the difference between the first two algal classes is equal or higher than 4% (cut-off threshold).

5.6.4.1. Emiliania huxleyi

This unicellular microalga is a well-known representative of the Coccolithophores. It is extensively studied as *E. huxelyi* is ubiquitous in marine environments. Its relative pigment composition as well as its growth rate are known to acclimate to different light irradiances.[164]
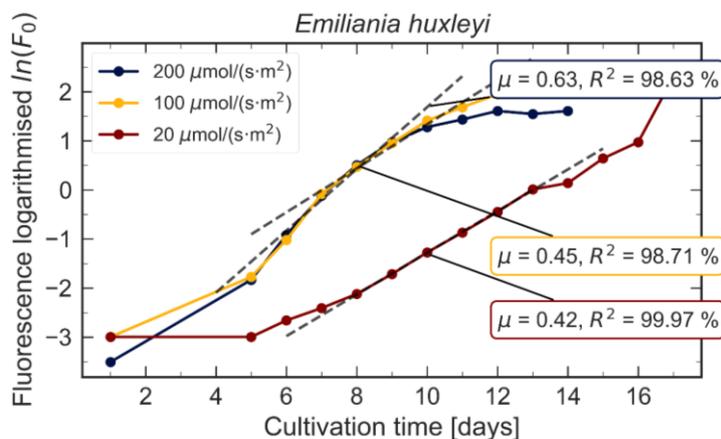
Figure S5-11: Growth rates under different light conditions of the haptophyte *E. huxleyi*. The growth rate μ is expressed in units per day.

Multiparametric Sensing – from Development to Application in Marine Science

Table S5-17 Growing conditions of *E. huxleyi* and days of examination after inoculation. The determined biomass at these days is given in cells per microliter.

growth phase	light condition (μmol / (s·m²))	cultivation time until examination	biomass (cells per microliter)	growth rate μ (1/d)
lag-phase	20	7	244	0.42
exponential phase		12	1,611	
stationary phase		14	1,863	
lag-phase	100	5	948	0.45
exponential phase		7	1,948	
stationary phase		12	11,625	
lag-phase	200	5	1,210	0.63
exponential phase		7	4,895	
stationary phase		14	14,050	

Table S5-18 Changes in the predicted group membership of the haptophyte E. huxleyi during growth, when cultivated under low light intensity (20 µmol / (s m2)). A statistically secure result is marked with *

lag-phase		exponential phase		stationary-phase	
identified phylum	probability (%)	identified phylum	probability (%)	identified phylum	probability (%)
haptophytes*	31.09%	dinoflagellates	26.27%	haptophytes	26.38%
diatoms	23.30%	haptophytes	23.65%	dinoflagellates	22.09%
chlorophytes	18.11%	ochrophytes	20.73%	ochrophytes	19.96%
dinoflagellates	17.33%	chlorophytes	9.03%	diatoms	11.81%
ochrophytes	7.25%	cryptophytes	9.00%	cryptophytes	9.15%
cryptophytes	2.88%	diatoms	8.90%	chlorophytes	8.63%
euglenophytes	1.05%	euglenophytes	3.42%	ruglenophytes	2.97%
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-19 Pigment : chlorophyll-*a* molar ratios of haptophyte *E. huxleyi* studied after growing under 20 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	Chl-c3	MgDVP	Chl-c2	Fuco	4KHF	19HF	Diadino	Diato	GyrE (12:0)	Np-c2	Chl-a	Chl-a'	βε-Car	ββ-Car
lag-phase	0.17	--	0.21	0.67	0.12	0.27	0.09	--	0.02	0.12	1.00	--	0.02	0.02
exponential	0.18	--	0.24	0.69	0.15	0.26	0.07	--	0.02	0.11	1.00	--	0.02	0.02
stationary	0.25	0.01	0.33	0.69	0.31	0.31	0.12	0.01	0.01	0.13	1.00	0.01	0.01	0.03

Table S5-20 Changes in the allocated group membership of the haptophyte *E. huxleyi* during growth, when cultivated under normal light intensity (100 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	probability (%)	identified phylum	probability (%)	identified phylum	probability (%)
chlorophytes	23.59%	haptophytes*	27.37%	haptophytes	29.35%
haptophytes	19.64%	dinoflagellates	22.65%	dinoflagellates	27.97%
ochrophytes	17.21%	ochrophytes	18.49%	ochrophytes	17.77%
diatoms	17.29%	diatoms	10.07%	chlorophytes	8.70%
dinoflagellates	13.52%	cryptophytes	9.30%	cryptophytes	7.36%
cryptophytes	7.04%	chlorophytes	8.70%	diatoms	7.26%
euglenophytes	2.71%	euglenophytes	3.42%	euglenophytes	3.60%
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-21 Pigment : chlorophyll-*a* molar ratios of haptophyte *E. huxleyi* studied after growing under 100 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	Chl-*c3*	[MV]-Chl-*c3*	MgDVP	Chl-*c2*	Fuco	4KHF	19HF	Diadino	Diato	GyrE-like	GyrE (12:0)	Np-*c2*	Chl-*a*	Chl-*a'*	$\beta\beta$-Car
lag-phase	0.13	--	0.02	0.31	0.34	0.08	0.72	0.12	0.04	--	0.02	0.11	1.00	--	0.03
exponential	0.16	--	0.01	0.30	0.36	0.10	0.63	0.07	0.07	--	0.01	0.13	1.00	0.01	0.04
stationary	0.20	0.01	0.01	0.33	0.34	0.19	0.79	0.24	0.08	0.01	0.01	0.13	1.00	0.01	0.04

Table S5-22 Changes in the predicted group membership of the haptophyte *E. huxleyi* during growth, when cultivated under high light intensity (200 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	probability (%)	identified phylum	probability (%)	identified phylum	probability (%)
chlorophytes*	24.89%	haptophytes*	27.61%	haptophytes	27.48%
haptophytes	19.00%	dinoflagellates	19.93%	dinoflagellates	27.16%
ochrophytes	17.77%	ochrophytes	17.49%	ochrophytes	18.31%
diatoms	17.56%	cryptophytes	10.76%	chlorophytes	10.83%
dinoflagellates	13.31%	chlorophytes	9.99%	cryptophytes	7.88%
cryptophytes	7.78%	diatoms	9.20%	diatoms	7.15%
euglenophytes	2.70%	euglenophytes	6.02%	euglenophytes	4.19%
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-23 Pigment : chlorophyll-*a* molar ratios of haptophyte *E. huxleyi* studied after growing under 200 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	Chl-*c3*	[MV]-Chl-*c3*	MgDVP	Chl-*c2*	Fuco	4KHF	19HF	Diadino	Diato	GyrE-like	GyrE (12:0)	Np-*c2*	Chl-*a*	Chl-*a'*	*βε*-Car	*ββ*-Car
lag-phase	0.12	–	0.02	0.31	0.27	0.06	0.80	0.14	0.12	–	0.02	0.10	1.00	0.03	–	0.04
exponential	0.14	–	0.01	0.28	0.27	0.09	0.65	0.10	0.15	–	0.01	0.11	1.00	–	–	0.04
stationary	0.15	0.05	0.01	0.27	0.15	0.17	0.96	0.44	0.09	0.02	0.01	0.12	1.00	–	0.01	0.03

5.6.4.2. Thalassiosira weissflogii

Conticribra weissflogii (= *Thalassiosira weissflogii*) is a bloom-forming diatom that can be found both in freshwater and marine environments. This diatom is extensively studied as it is known that its pigment composition is sensitive to different light conditions and is adjusted during growth.[165]
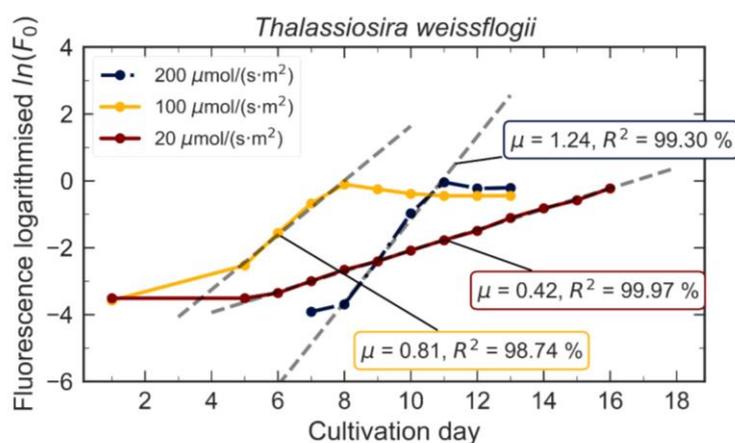
Figure S5-12: Growth rates under different light conditions of the diatom *T. weissflogii*. The growth rate μ is expressed in units per day.

Table S5-24 Growing conditions of *T. weissflogii* and days of examination after inoculation. The determined biomass at these days is given in cells per microliter.

growth phase	light condition (μmol / (s·m²))	cultivation time until examination	biomass (cells per microliter)	growth rate μ (1/d)
lag-phase	20	10	10	0.42
exponential phase		16	60	
Ssationary phase		18	145	
lag-phase	100	7	71	0.81
exponential phase		9	80	
stationary phase		12	66	
lag-phase	200	4	41	1.24
exponential phase		7	102	
stationary phase		10	129	

Table S5-25 Changes in the allocated group membership of the diatom *T. weissflogii* during growth, when cultivated under low light intensity (20 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	probability (%)	identified phylum	identified phylum	probability (%)	probability (%)
dinoflagellates	47.42%	diatoms	22.80%	diatoms*	26.62%
cryptophytes	43.68%	haptophytes	19.61%	haptophytes	21.12%
diatoms	11.56%	cryptophytes	18.84%	cryptophytes	17.08%
ochrophytes	10.26%	ochrophytes	17.35%	ochrophytes	16.70%
haptophytes	7.17%	dinoflagellates	10.58%	chlorophytes	11.42%
euglenophytes	6.86%	chlorophytes	8.85%	dinoflagellates	7.27%
chlorophytes	6.05%	euglenophytes	2.96%	euglenophytes	2.795
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-26 Pigment : chlorophyll-*a* molar ratios of diatom T. weissflogii studied after growing under 20 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	Chl-*c2*	Chl-*c1*	Fuco	Diadino	Diato	Chl-*a*	Chl-*a'*	$\beta\beta$-Car
exponential	0.04	0.03	0.39	0.09	0.04	1.00	0.02	0.04
late-exponential	0.06	0.07	0.52	0.16	0.03	1.00	--	0.05
stationary	0.05	0.05	0.47	0.14	0.02	1.00	--	0.05

Table S5-27 Changes in the allocated group membership of the diatom *T. weissflogii* during growth, when cultivated under normal light intensity (100 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
probability (%)	identified phylum	identified phylum	probability (%)	identified phylum	probability (%)
diatoms	23.13%	diatoms*	32.91%	cryptophytes*	30.51%
ochrophytes	20.12%	haptophytes	19.37%	diatoms	22.14%
haptophytes	19.83%	chlorophytes	17.40%	ochrophytes	17.67%
chlorophytes	13.54%	dinoflagellates	12.13%	haptophytes	11.79%
cryptophytes	13.19%	ochrophytes	9.11%	dinoflagellates	7.76%
dinoflagellates	7.03%	cryptophytes	8.37%	chlorophytes	7.76%
euglenophytes	4.16%	euglenophytes	1.71%	euglenophytes	3.37%
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-28 Pigment : chlorophyll-*a* molar ratios of diatom *T. weissflogii* studied after growing under 100 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	MgDVP	Chl-*c2*	Chl-*c1*	Fuco	Diadino	Diato	Chl-*a* allomer	Chl-*a*	Chl-*a'*	*ββ*-Car
exponential	0.01	0.04	0.03	0.38	0.18	0.01	0.01	1.00	0.01	0.05
late-exponential	--	0.06	0.04	0.41	0.24	0.13	0.01	1.00	0.02	0.04
stationary	--	0.06	0.03	0.42	0.47	0.07	--	1.00	0.02	0.06

Table S5-29 Changes in the predicted group membership of the diatom *T. weissflogii* during growth, when cultivated under high light intensity (200 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	identified phylum	probability (%)	identified phylum	identified phylum	probability (%)
ochrophytes	22.29%	ochrophytes*	24.82%	ochrophytes	20.96%
diatoms	20.08%	euglenophytes	16.60%	diatoms	17.01%
haptophytes	17.63%	cryptophytes	14.89%	cryptophytes	17.68%
chlorophytes	17.19%	diatoms	13.77%	chlorophytes	16.18%
cryptophytes	12.30%	chlorophytes	12.37%	euglenophytes	13.65%
dinoflagellates	6.78%	dinoflagellates	10.97%	haptophytes	10.45%
euglenophytes	6.73%	haptophytes	7.59%	dinoflagellates	7.08%
cyanobacteria	0.00%	cyanobacteria	0.00%	cyanobacteria	0.00%

Table S5-30 Pigment : chlorophyll-*a* molar ratios of diatom *T. weissflogii* studied after growing under 200 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state).

	Chlyde *a*	MgDVP	Chl-*c2*	Chl-*c1*	Fuco	Diadino	Diato	Chl-*a* allomer	Chl-*a*	Chl-*a'*	*ββ*-Car
Exponential	--	0.01	0.03	0.02	0.39	0.30	0.01	0.02	1.00	0.02	0.06
Late-exponential	--	--	0.04	0.03	0.40	0.41	0.10	--	1.00	0.02	0.06
Stationary	0.03	--	0.05	0.04	0.40	0.51	0.19	0.01	1.00	0.03	0.09

5.6.4.3. Synechococcus sp.

This widely distributed picocyanobacteria occur in marine as well as in freshwater environments and encompass a high pigment variety. Its relative pigmentation is sensitive to given environmental conditions.[166,167]
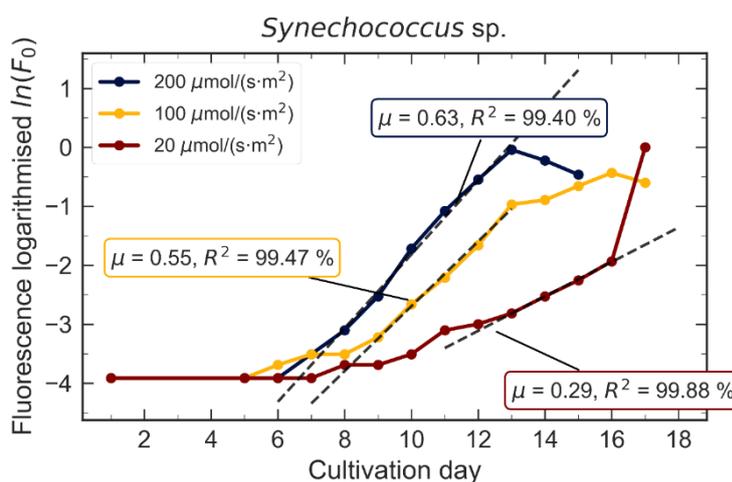
Figure S5-13: Growth rates under different light conditions of the cyanobacteria *Synechococcus* sp.. The growth rate *μ* is expressed in units per day.

Table S5-31 Growing conditions of *Synechococcus* sp. and days of examination after inoculation. The determined biomass at these days is given in cells per microliter.

growth phase	light condition (μmol / (s·m²))	cultivation time until examination	biomass (cells per microliter)	growth rate μ (1/d)
lag-phase	20	10	2,500	0.29
exponential phase		12	19,480	
stationary phase		15	986,000	
lag-phase	100	9	75,300	0.55
exponential phase		12	101,570	
stationary phase		15	323,000	
lag-phase	200	8	2,344	0.63
exponential phase		10	11,650	
stationary phase		13	$604 \cdot 10^9$	

Table S5-32 Changes in the allocated group membership of the cyanobacteria *Synechococcus* sp. during growth, when cultivated under low light intensity (20 μmol / (s m²)). A statistically secure result is marked with ***.**

lag-phase		exponential phase		stationary-phase	
identified phylum	identified phylum	probability (%)	identified phylum	identified phylum	probability (%)
cyanobacteria*	98.48%	cyanobacteria*	99.98%	cyanobacteria*	99.98%
dinoflagellates	1.52%	dinoflagellates	0.02%	dinoflagellates	0.02%
haptophytes	0.00%	haptophytes	0.00%	haptophytes	0.00%
cryptophytes	0.00%	cryptophytes	0.00%	cryptophytes	0.00%
ochrophytes	0.00%	ochrophytes	0.00%	ochrophytes	0.00%
diatoms	0.00%	diatoms	0.00%	diatoms	0.00%
chlorophytes	0.00%	chlorophytes	0.00%	chlorophytes	0.00%
euglenophytes	0.00%	euglenophytes	0.00%	euglenophytes	0.00%

Table S5-33 Pigment : chlorophyll-*a* molar ratios of cyanobacteria *Synechococcus* sp. studied after growing under 20 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state). It was not possible to analyze the algae in lag-phase. However, phycobilines cannot be analyzed by the conducted HPLC method.

	Zea	Chl-*a*	Chl-*a*'	$\beta\beta$-Car
lag-phase*	--	--	--	--
exponential	1.96	1.00	0.02	0.26
stationary	2.51	1.00	--	0.64

Table S5-34 Changes in the predicted group membership of the cyanobacteria *Synechococcus* sp. during growth, when cultivated under normal light intensity (100 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	identified phylum	probability (%)	identified phylum	identified phylum	probability (%)
cyanobacteria*	99.62%	cyanobacteria*	99.98%	cyanobacteria*	99.96%
dinoflagellates	0.38%	dinoflagellates	0.02%	dinoflagellates	0.04%
haptophytes	0.00%	haptophytes	0.00%	haptophytes	0.00%
cryptophytes	0.00%	cryptophytes	0.00%	cryptophytes	0.00%
ochrophytes	0.00%	ochrophytes	0.00%	ochrophytes	0.00%
diatoms	0.00%	diatoms	0.00%	diatoms	0.00%
chlorophytes	0.00%	chlorophytes	0.00%	chlorophytes	0.00%
euglenophytes	0.00%	euglenophytes	0.00%	euglenophytes	0.00%

Table S5-35 Pigment : chlorophyll-*a* molar ratios of cyanobacteria *Synechococcus* sp. studied after growing under 100 μmol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state). However, phycoblines cannot be analyzed by the conducted HPLC analysis.

	Zea	Chl-*a*	Chl-*a'*	$\beta\beta$-Car
Lag-phase	1.39	1.00	--	0.49
Exponential	1.71	1.00	--	0.52
Stationary	0.90	1.00	0.02	0.29

Table S5-36 Changes in the predicted group membership of the cyanobacteria *Synechococcus* sp. during growth, when cultivated under high light intensity (200 μmol / (s m²)). A statistically secure result is marked with *.

lag-phase		exponential phase		stationary-phase	
identified phylum	identified phylum	probability (%)	identified phylum	identified phylum	probability (%)
cyanobacteria*	99.76%	cyanobacteria*	99.96%	cyanobacteria*	100.00%
dinoflagellates	0.24%	dinoflagellates	0.04%	dinoflagellates	0.00%
haptophytes	0.00%	haptophytes	0.00%	haptophytes	0.00%
cryptophytes	0.00%	cryptophytes	0.00%	cryptophytes	0.00%
ochrophytes	0.00%	ochrophytes	0.00%	ochrophytes	0.00%
diatoms	0.00%	diatoms	0.00%	diatoms	0.00%
chlorophytes	0.00%	chlorophytes	0.00%	chlorophytes	0.00%
euglenophytes	0.00%	euglenophytes	0.00%	euglenophytes	0.00%

Table S5-37 Pigment : chlorophyll-*a* molar ratios of cyanobacteria *Synechococcus* sp. studied after growing under $200\,\mu$mol / (s m²) and analyzed in three different growth phases (lag-phase, exponential phase and steady state). However, phycobilines cannot be analyzed by the conducted method.

	Zea	Chl-*a*	$\beta\beta$-Car
lag-phase	2.26	1.00	0.59
exponential	2.89	1.00	0.59
stationary	4.52	1.00	0.61

6. Graphical user Interface for the Evaluation of the Fluorescence Pattern of Algae

For data evaluation, a software algorithm based on the linear discriminant analysis (LDA) was written in Python. This pattern recognition analysis compares measurement information from an unknown sample to an established trainings database of known group affiliation. The prediction accuracy of the algorithm will improve with the number of training objects.

In this chapter, the operating introductions of the software interface, also referred to as graphical user interface (GUI), are presented, whereas the theoretical background is explained earlier in section 2.3.

6.1. Software Installation and Directory Structure

The data evaluation software is written in Python and an executable version of the Software is provided in the Appendix 11. Start the *ALPACAGUI.exe* program file by double-clicking.

It is advised to maintain the structure of the sub-folders as it is initially given, as the executable referrers to these folders. The directory structure consists of two folders – one for the measurement (in the following called '*measurement'*) and the other for supplementary files (in the following called '*supplementary'*), such as training database and calibration files. In the supplementary folder create three folders called '*calibration', 'phytoplankton'* and '*trainingdatabase'*.

In the '*calibration'* folder, all calibration measurements and parameters are stored. It is further recommended to separate the emission and excitation calibrations. In the '*phytoplankton'* folder a *.txt* document containing all taxonomic details are stored. Further a list linking the taxonomic level with a color information, has to be stored in this folder. In the folder '*trainingdatabase' .txt* files are stored containing the average fluorescence intensities of the training objects. These training objects should also be corrected on their emission and excitation side as well as according to their baseline.
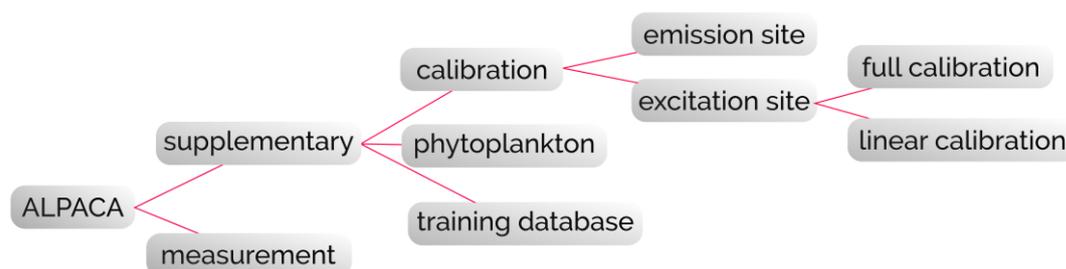
Figure 6-1: Recommended directory structure for the ALPACA software interface

6.2. Software Interface

After installing and starting of the software, the interface (see Figure 6-2) will appear on the screen.
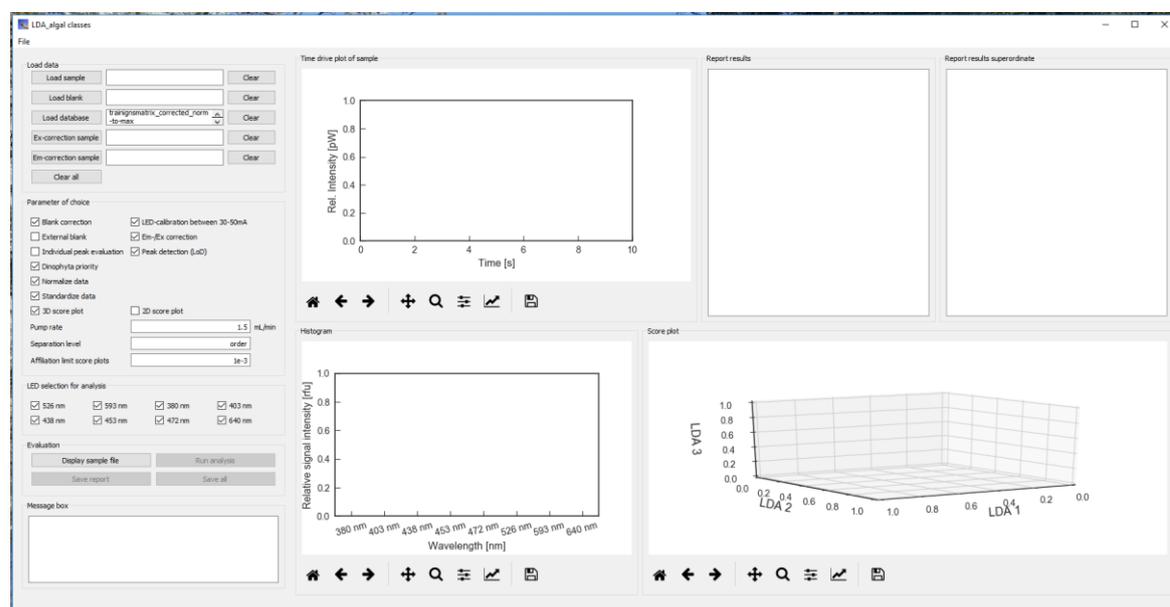
Figure 6-2: Interface of the data evaluation software to determine the group membership of the sample.

6.2.1. Load Data

On the left side on the top, you will find a section (*Load data*) to load different files required for evaluation. Besides the sample file itself, you can load an external blank file, a correction file to balance differences between different emission channels or between different LEDs.

When pressing the bottoms '**Load sample**' and '**Load blank**', a second window will appear on screen directed to the subfolder '*measurement*', which can be found in the same folder as the python script. While the sample file is mandatory, you do not need a blank file if you process a baseline correction. The next bottom '**Load database**' will open a sub-subfolder '*trainingdatabase*' in the subfolder '*supplementary*', which can be found in the same branch as the python script.

Further, correction parameters for the emission and excitation side have to be selected. To balance differences between the absolute emission intensity of the LEDs, you have to calibrate the ALPACA on the excitation side (see section 2.3.3.2). This correction is required as LEDs of the same fabrication might differ in their absolute peak wavelength as well as their peak wavelength is temperature dependent. These correction parameters saved in a txt-document, must be loaded by '**Ex-correction sample**'. Differences between the four photodiodes have to be corrected as well. The correction parameters saved in an txt-document can be loaded via '**Em-correction sample**'.

If you want to change or delete a file, press either '**Clear**' for a single selection or '**Clear all'** to delete each entry.

Figure 6-3: Screenshot of the section 'Load data' on the top left of the interface.

6.2.2. Training Database

For the software, the training database is fixed in a txt document to save process running time. The training objects should be pretreated similar to the sample, e.g. they should be normalized and/or standardized. Information about the pretreatment should be stored in the name of the txt file.

In the subfolder '*supplementary/trainingdatabase'*, the user can find some examples for the training database. The name indicates whether rhodophyta or cyanobacteria are part of the training database or not or whether the training objects are autoscaled (corrected) or not.

When the user wants to update the training database, a hardcoded version of the training database has to be constructed first, which can be used in the software interface. For this purpose, use the python script '*creating_trainingdatabase.ipynb'*. This script can be used with an Integrated Development Environment (IDE).

At the end of the script (line 1468), the user defines the directory to the folder containing the training objects (training_db). Normally, the training objects are stored in a subfolder under the directory '*supplementary/trainingdatabase/trainingobjects'*. The training objects were measured with the ALPACA and saved under the filename '*DateTime_sample-ID_LEDintensity_MAZeT-amplification.txt'*. As sample-ID use the algae family name without species information. The sample-ID should match with the information provided in the phytoplankton folder (*'supplementary/phytoplankton'*). The directory to these information is described in the next step. This folder contains information about the phytoplankton and its taxonomic classification. These txt-documents have to be updated according to the training objects.

Between line 1473-1482, the user can define some process parameters which are needed for evaluation. For instance, the user has to define at which taxonomic level the separation has to be conducted and whether the training objects should be corrected, normalized and/or standardized.

In the next paragraph (line 1485-1492), the user defines, which LEDs were used during the measurement.

When all process parameters are defined, the script can be executed and the training database is constructed and saved in the folder '*supplementary/trainingdatabase*'.

6.2.3. Parameters of Choice

After uploading the sample measurement into the software, calibration and standardization parameters can be defined in the software interface. Further parameters relevant for the evaluation can be selected by the user, e.g. whether the algorithm should focus on the distinct identification of dinoflagellates or not. Some parameters are selected by default. Besides the selection of some parameters, you can also define some parameters, for example the pump rate (mL/min) which was used during measurement.

Figure 6-4: Screenshot of the section 'Parameter of choice'.

Blank correction: meaning the subtraction of noise or biofouling which does not derive from the sample itself. As phytoplankton can grow in the capillary, when the capillary is insufficiently cleaned, a blank solution should be measured each time right before the sample is analyzed. This parameter is selected by default.

External blank: if the blank values are not saved in the sample file for some reasons (the software was broken down and a blank value couldn't be saved before or the blank measurement was not conducted before the sample measurement). If this optional parameter is selected, you have to define a blank file in the section above ('**Load blank**').

LED-calibration between 30-50 mA: normally, the amplification of the LEDs should be set between 30-50 mA. In this range, the amplification operates linearly. However, it is possible to

measure at lower or higher amplifications. If this is the case and you prefer to measure outside the linear range, clear this checkbox.

Em-/Ex-correction: If this parameter is selected (default), the sample will be corrected on the emission as well as on the excitation side.

Peak detection (LoD): If this checkbox is selected, the average fluorescence intensity is corrected by the average noise signal. Only signals higher than the detection limit ($\bar{x}_{\text{blank}} + 3 \cdot \delta_{\text{blank}}$) are considered for evaluation. This checkbox is selected by default.

Dinophyta priority: The LDA can be trained to focus on the separation of dinoflagellates from other algae. If this checkbox is selected (default value), the identification of dinoflagellates is more reliable than when the checkbox is cleared meaning that the false-positive rate increases slightly while the false-negative rate decreases.

Normalize data and **Standardize data:** For autoscaling of the sample and training objects. Normally, the data matrix depends on the scaling and the neutral point of the variables. This dependence biases the classification results leading to wrong group allocations. Therefore, all samples should be autoscaled before evaluation (default).

3D score plot or **2D score plot:** One can decide whether the results should be presented in a two dimensional or a three dimensional plot. A 3D plot contains more information about the group separation than the 2D plot. However, a 3D plot can be confusing due to the overlap of several algal groups.

Pump rate: It is advised to measure with a pump rate of 1.5 mL/min or below. This flow rate ensures that each cell event is recognized at each measurement channel with at least three measurement points. However, the flow rate has to be adapted according to the conducted experiment.

Separation level: The user can select the taxonomic level at which the sample is separated. The user can select between '*phylum*', '*class*', '*order*' or '*family*'. The appropriate choice depends on the training database. The default entry is '*order*' as it was determined, that the figures of merit are most suitable at this taxonomic level.

6.2.4. LED Selection for Analysis

The ALPACA features eight excitation channels enabling the recording of the emitted fluorescence intensity upon excitation at eight different excitation wavelengths. The ALPACA analyzes the relative contribution of the different excitation wavelengths to the fluorescence intensity emitted from the biomass in the capillary. This contribution to the resulting fluorescence intensity is directly linked to the relative pigment pattern within the photosystems as it was explained earlier (compare section 2.3.2). If pigments are present in the photosystems that absorb in same wavelength range

where the LED emits, the emitted fluorescence intensity is higher than without this pigment contribution. For example, if peridinin is present in the photosystem, this can be excited between 470-530 nm. When using an LED in this wavelength range, peridinin is excited. The absorbed energy is guided to the reaction center and surplus energy (that cannot be used for photosynthesis) is emitted from the reaction center. Consequently, the higher the amount of absorbed energy, the higher the emitted fluorescence intensity.

During the development of the ALPACA, we found that the following LEDs are most suitable for separation and classification of dinoflagellates and cyanobacteria from other algal groups:

- 380 nm
- 403 nm

- 438 nm
- 453 nm

- 472 nm
- 526 nm

- 593 nm
- 640 nm

However, if a separation analysis with less LEDs is desired, the user can select in the next section, which LEDs should be included for evaluation.
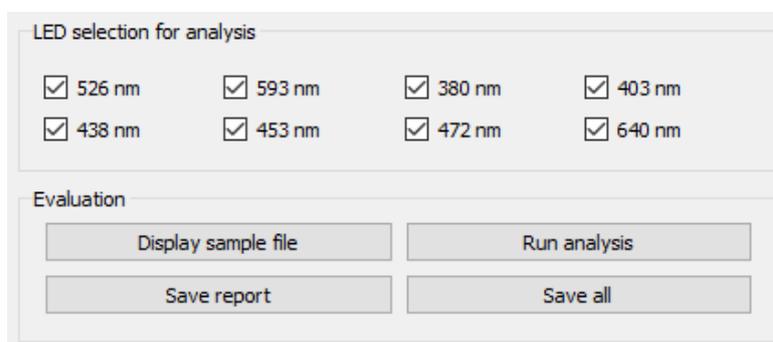
Figure 6-5: Screenshot of the section 'LED selection' and 'Evaluation'.

6.2.5. Evaluation

After loading the required txt-documents and defining all process parameters, press the bottom '**Display sample file**'. If the requirements are fulfilled, the time drive plot is loaded displaying the recorded fluorescence intensity over time upon excitation at different excitation wavelengths. When the sample measurement is loaded and the sample region as well as the baseline is defined, press '**Run analysis**' and the linear discriminant analysis (LDA) is processed with the parameters chosen in advance. The output of the LDA is displayed in the histogram (see section 6.2.6.2), the score plot (see section 6.2.6.3) and the result tables (see section 6.2.6.4).

If only the result tables (containing the assigned group affiliation) should be saved, press '**Save**'. A folder called '*results_LDA_...*' is created. The name of the folder contains all chosen parameters, such as blank (for blank correction), correction (for emission and excitation correction) and/or peak (for peak detection). A second folder '*scores*' is created containing the scores of the sample and the

training objects. If the histogram and the score plots should be saved as well, press '**Save all**'. A folder '*LDA_figures*' is created in the same directory as the measurement folder.

6.2.6. Plotting Areas

On the right side of the interface, several plotting areas are defined. Besides the figure on the top where the time-drive plot of the sample is displayed, a histogram and a 2D or 3D score plot area is shown on the bottom.

6.2.6.1. Time Drive Plot of Sample

After loading the measurement, the fluorescence intensity, which was recorded over time at up to eight different excitation channels, is displayed in the time-drive plot (Figure 6-6).
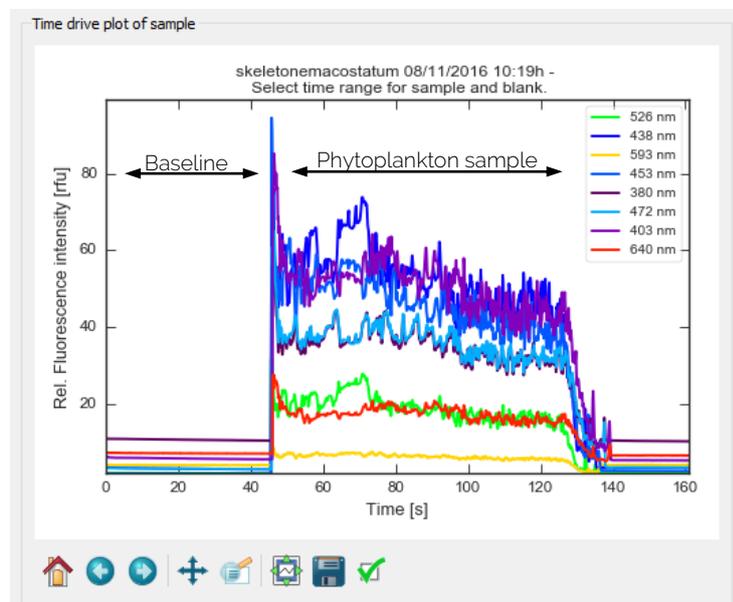
Figure 6-6: Time-drive plot displaying the emitted fluorescence intensity over time upon excitation at different wavelengths. The different parts of the measurement are marked and explained in the plot.

First, in this time-drive plot, the user has to mark the sample region which has to be evaluated and then the baseline for correction has to be defined. This is done by pressing '**Strg**' and, simultaneously, selecting the start and stop point (for the sample and the baseline) in the time drive plot using the cursor.
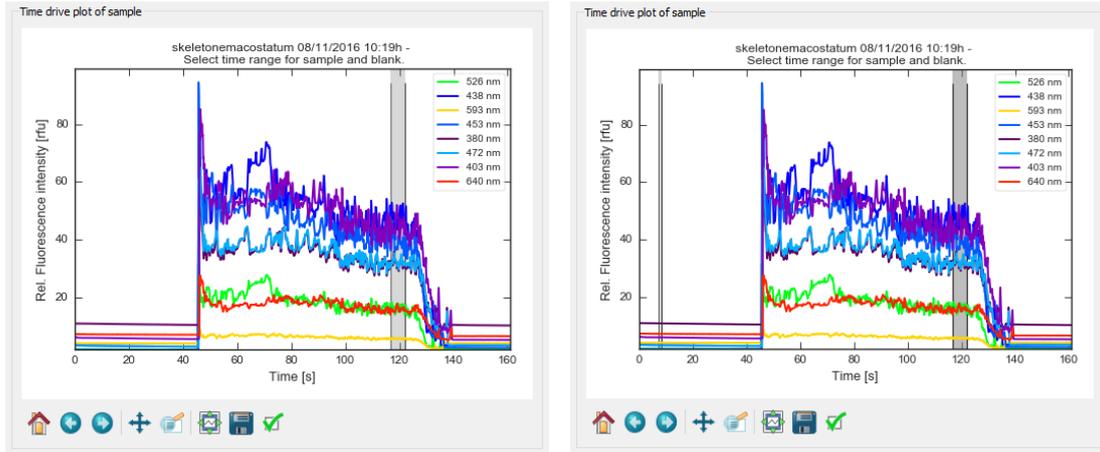
Figure 6-7: Time-drive plot displaying the emitted fluorescence intensity over time upon excitation at different wavelengths. Left: the start and stop point of the sample is selected. Afterwards, the baseline for correction is defined as well, in the right plot.

6.2.6.2. Histogram

When applying the LDA, the average fluorescence intensity is calculated. The average intensity is corrected, if previously defined and displayed as histogram in the plot on the bottom-left side. This histogram is described as relative fluorescence pattern, describing an excerpt of a whole excitation spectrum, normally measured with a fluorescence spectrofluorometer.
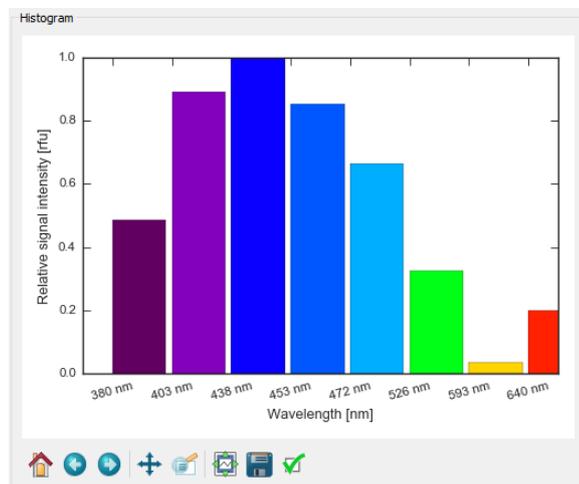
Figure 6-8: Fluorescence pattern of the sample illustrating the average and corrected contribution of different pigments to the resulting fluorescence intensity recorded on the ALPACA.

6.2.6.3. Score Plot

After applying the LDA, training objects belonging to the same algal group, are displayed as centroids in the score plot. The centrum of the centroids is associated with the group median and the solid lines defining the algal centroids correlate with the first, second and third standard

deviation in each projection direction. The score of the sample is displayed in the score plot as orange triangle.
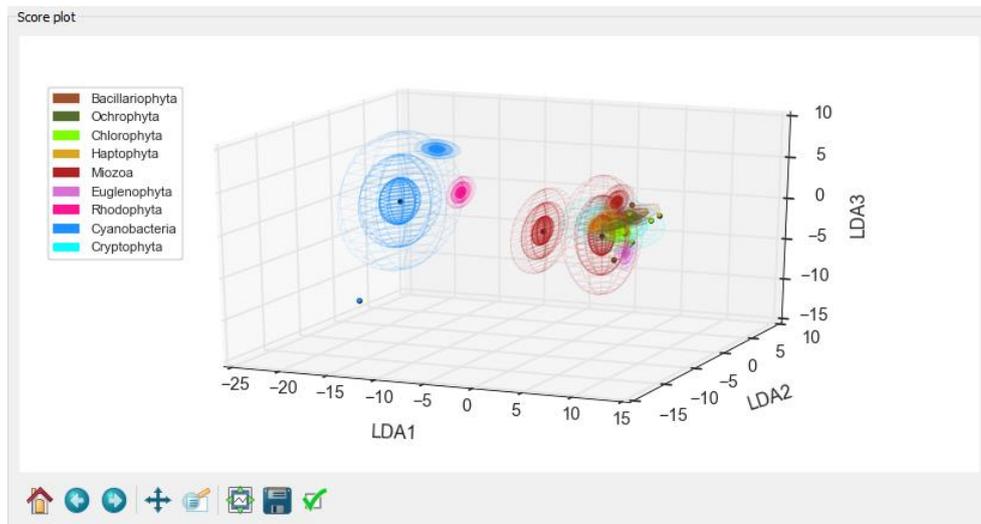
Figure 6-9: 3D score plot of the training objects displayed as centroids. The sample is illustrated as orange triangle, but due to the great spectral overlap of the algal centroids, the sample cannot be seen in this plot.

6.2.6.4. Message Box

The message box is located on the bottom on the left side. At this time, the massage box is more for the developer than for the user as it informs if the algorithm works correct and if the saving was processed.

6.2.6.5. Report

The report tables are located on the right side on the top of the interface. The left report table displays the assigned group affiliation as it was defined by the user; e.g. if the user has chosen a separation at the '*order*' level, the identified affiliation is described at the order level. Further, the probability is related to the Mahalanobis distance and describes the absolute probability of group membership. By contrast, the relative probability is displayed in the right table. Here, the affiliation probability sums up to 100%.

7. Additional Results and Field Trips

7.1. Initial Stability Test in the Port of Genoa

The first field trip within the SCHeMA-project was conducted in the port of Genoa in Italy between January 18th and March 7th, 2017. All project partners participated in this field trip which was dedicated to the general integration of all sensors into one shared submersible housing. After successfully integration of the instrumental electronics, a long-term experiment was performed during mid-February and beginning of March. During this period of time, the ALPACA analyzed water samples directly on-site and investigated the occurrence of algae or cyanobacteria in the harbor. Although the phytoplankton concentration was expected to be very low, as algal blooms normally arise during spring and autumn. This field trip was further a performance test concerning the detection limit of the ALPACA. Results of this study were published as first real-world application of the ALPACA in the corresponding manuscript (compare chapter 4).

Besides the performance of the ALPACA, the long-term stability was evaluated under harsh conditions in natural environments. Hereby, one focus lies on possible electromagnetic interferences and couplings from other sensors enclosed in the same submersible housing. Initial tests confirmed that the ALPACA suffers significantly from those interferences leading to a high signal noise level as it is shown in Figure 7-1. Therefore, it was not possible to analyze individual cell signals as the average signal intensity of individual cells is much lower with only few nanowatts. After introduction of an additional grounding, the noise level has been reduced by a factor of ten to the initial background level as it is shown in Table 7-1.
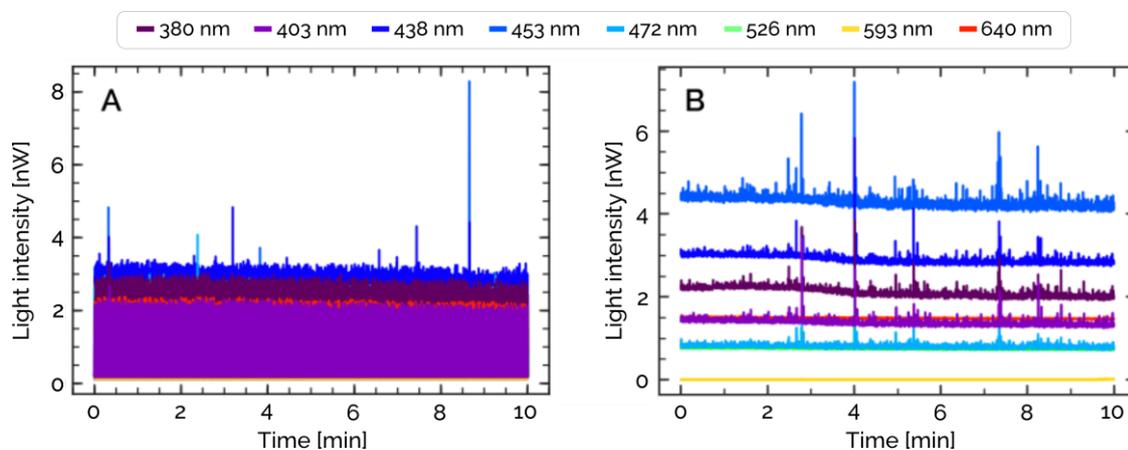
Figure 7-1: Average signal level before and after grounding recorded during the preparation of the first field trip at the port of Genoa. Before the additional grounding was implemented, the signal noise was too high to identify individual cell signals **(A)**, whereas, the average noise level has been reduced after grounding by a factor of ten compared to the initial background level **(B)**.

Table 7-1 Average standard deviation of the signal before and after additional grounding recorded during the preparation of the field trip at the port of Genoa. Additional grounding reduced the signal noise by a factor of ten.

	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
Before	0.85	0.64	0.93	0.88	0.80	0.76	0.76	0.66
After	0.08	0.09	0.07	0.12	0.04	0.02	0.01	0.02

To summarize, the initial field trip in the port of Genoa has proven that the ALPACA must be carefully grounded, in particular when the system is combined with other electrical components in one submersible housing. Moreover, the material of choice must be reconsidered, as the initially chosen aluminum foil has corroded the electrical components, when it came into contact with the sea water.

7.2. Arcachon Bay

This field trip was conducted between May 15th and May 19th, 2017 in the Bay of Arcachon, France, and it mainly targeted the validation of the ALPACA in coastal environments by commercially available in-situ fluorometers. During that time, both device versions of the ALPACA, the mobile one and the submersible one, were used for evaluation. While the mobile device can be used on-site for direct evaluation of measurement data, the submersible version can be integrated into a shared submersible housing for joint studies of multiple sensing systems.

Figure 7-2: Sampling sites during the field trip in Arcachon Bay from May 15 to 19, 2017. The geophysical cycle of 24 hours (marked in red) was conducted in Comprian, whereas the night cycle was performed in the harbor of Arcachon (marked in yellow). Furthermore, samples from different sampling sites in the southern part of the Bay were compared to each other (G1-G3), following a pollution gradient towards the industrial zone of Arcachon in Gujan-Mestras.

For validation purposes, four different experiments were conducted at a temporal and spatial scale: First, the average biomass and alterations in the algal population were investigated during a geophysical diurnal cycle (24 hours) in the Bay of Arcachon at Comprian (compare Figure 7-2). During that time, samples were taken at hourly intervals and the preliminary evaluation, providing rapid information whether phytoplankton are present in the sample, were analyzed on-site. A detailed evaluation of the average algal composition was examined afterwards using the graphical user interface for LDA as presented in chapter 6. Second, the stand-alone device was integrated into a submersible housing and deployed in the harbor of Arcachon for continuous measurement overnight, studying the qualitative and quantitative differences between Comprian and the harbor.

Third, alterations of the algal population were investigated along a pollution gradient in front of the industrial zone of Arcachon. In the last experiment, samples from different sampling sites owned by the French observation center (SOMLIT), were analyzed.

For validation, either a TriLux algae detector from Chelsea (Chelsea Technologies Group Ltd, chelsea.co.uk) or a Seapoint Chlorophyll Fluorometer (seapoint.com) was used for quantification. The chlorophyll fluorometer was equipped on a CTD probe. The relative algal composition was validated by the University of Genoa following the Utermöhl method described by Hasle and Zingone et al.[141,142]

Based on the knowledge gained from the previous field trip, the connectors linking the fluidic system from outside with the capillary inside the optics block of the ALPACA, were updated and the standard barbed fluidic connectors were replaced. These connectors are now made out of stainless steel. Stainless steel improves the grounding of the ALPACA as well as it prevents the system against electromagnetic interferences, in order to reduce the average background level as well as signal noise to a minimum. In Table 7-2, the average blank level over time is shown for both device versions. Repetitive controls of the bank level were suitable, in particular for the submersible device, to control whether the implemented cleaning steps were sufficient or whether biofouling occurred during the field trip. For cleaning purposes, 70% ethanol was used as common disinfectant. The average blank level increased slightly over time indicating that the cleaning solution ought to be revised and preservatives, such as sodium azide, might be useful in subsequent field trips.

Table 7-2 Average blank level and its standard deviation for all LEDs as comparison between the first and the last experiment. [a]

Mobile APACA	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
1st experiment mean	168.7	530.7	1187.7	821.3	53.4	50.1	0.23	6984.8
1st experiment deviation	33.5	26.2	97.7	49.2	23.8	16.4	0.4	63.2
2nd experiment mean	790.2	594.8	1004.2	536.5	272.7	227.4	0.27	927.7
2nd experiment deviation	30.2	22.5	29.7	18.3	12.6	5.3	0.4	8.0
Standalone APACA	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
1st experiment mean	1067.27	706.23	1296.57	687.79	441.43	322.4	59.82	1127.87
1st experiment deviation	504.71	194.8	514.86	273.14	320.06	185.97	127.03	409.37
2nd experiment mean	789.66	547.11	918.63	497.69	261.19	191.14	0.36	781.19
2nd experiment deviation	350.1	243.07	407.52	219.08	116.56	84.4	0.6	345.2

[a] Values are given in light intensity in picowatts.

7.2.1. Geophysical Cycle at Comprian

For the geophysical diurnal cycle, Comprian has been selected as appropriate sampling site. The mobile version of the ALPACA was installed on board and surface water samples were taken at a

depth of 2.5 meters under the surface on an hourly base. Fluctuations in the measurement level arise due to the tidal currents. For the analysis, 10 mL of sample were pumped through the ALPACA and the relative fluorescence pattern was determined on-site (three-fold determination). Further, variations of the total biomass concentration were examined and validated against the TriLux algae detector from Chelsea Technologies Ltd. Results of this validation are shown in Figure 7-3A-B. Apart from few outliers, the overall trend of both LEDs targeting either the detection of algae (Figure 7-3A) or the detection of cyanobacteria (Figure 7-3B), coincides with the reference sensor from Chelsea. The overall trends of the biomass content are further confirmed by the turbidity measurement (Figure 7-3C). Furthermore, these alterations of the biomass concentration are compared with the oxygen saturation, recorded by the integrated CTD sensor from IDRONAUT. The oxygen saturation follows the general principles of photosynthesis and respiration during day and night (Figure 7-3D). As the initial step of the photosynthesis is light dependent, photosynthesis only takes place during day while respiration takes place during night. During these hours of darkness, oxygen is consumed by the phytoplankton and simultaneously, the biomass increased in the surface water, due to fluctuations caused by the tidal currents. Therefore, the effect of oxygen consumption due to algal respiration is intensified. By contrast, oxygen is produced by algae due to photosynthesis during the day leading to an oxygen saturation of more than 100%.

Besides the semi-quantification of phytoplankton, the average algal composition was examined using the LDA. These results have been randomly tested by the microscopic evaluation conducted by the University of Genoa. For simplicity reasons, measurement results are presented in Figure 7-4, Table 7-3 and Table 7-4 at 4-hour intervals. The mobile and the submersible system, both agree that diatoms and ochrophytes are the dominant groups, while cyanobacteria are absent in this experiment. However, the mean values for the average algal composition differs widely between the microscopic evaluation and the ALPACA and, in particular, the standard deviation for the ALPACA is high. These discrepancies in the results might originate from the different sampling processing. For microscopic evaluation, only small amounts of sample are taken on a random bias, whereas the ALPACA analyzes 10 mL of sample and calculates the average sample composition according to the algal groups included into the training database. Therefore, trends can be compared between the microscopic analysis and the ALPACA; although absolute values are hardly comparable.
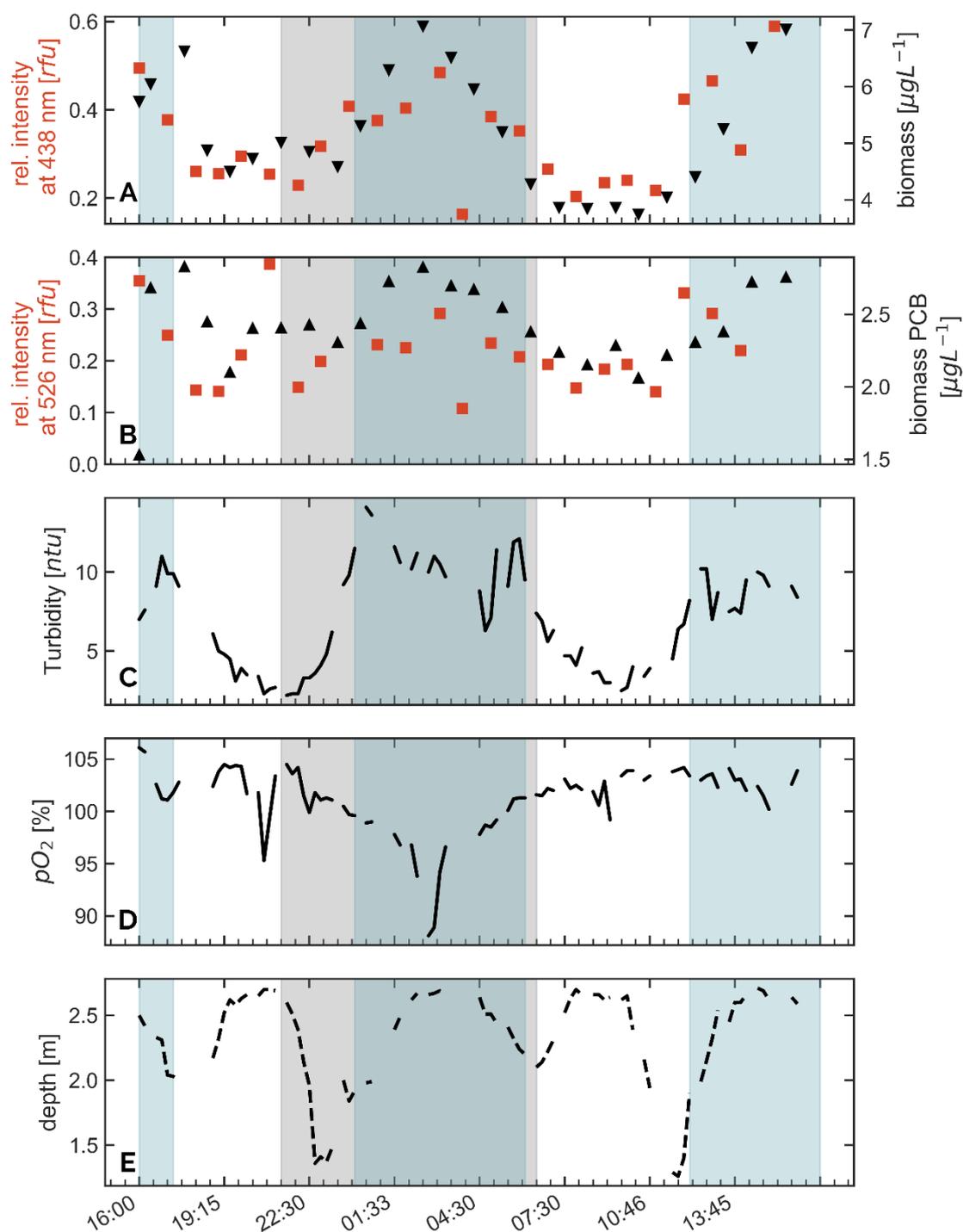
Figure 7-3: Relative fluorescence intensity recorded by the ALPACA upon excitation at either 438 nm or 526 nm targeting the detection of chlorophyll-*a* (**A**) or phycoerythrin and phycocyanin as diagnostic pigments for rhodophytes and cyanobacteria. (**B**). In each case, the relative fluorescence intensity is compared to the TriLux algae sensor from Chelsea Technologies (marked in black). In case of the cyanobacteria, the reference sensor targets the detection of phycocyanobilins (PCB), which are found in cyanobacteria and rhodophytes. Besides, the turbidity (**C**), the oxygen saturation (**D**) and the average depth (**E**), where the samples were taken, are presented. Trends for turbidity and biomass coincide, while the trend of the oxygen saturation follows the general principle of photosynthesis and respiration during day and night. Changes in the measurement level (**E**) reflect the tides over day. Hours of darkness are marked in grey, whereas low tides are marked in light blue.
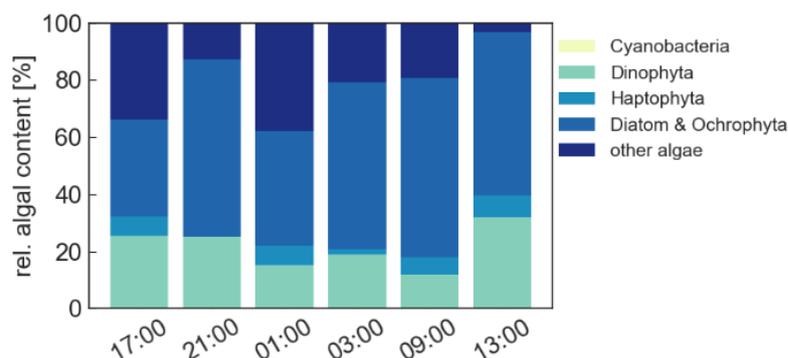
Figure 7-4: Alterations in the relative algal composition during the geophysical circle. At no time, cyanobacteria have been detected by the algae detection unit.

Table 7-3 Relative algal composition and its variations during the geophysical cycle examined by the algae detection module. Relevant phytoplankton groups are marked in bold letters.

	daylight		night		daylight		Mean [%]	SD [%]
	17:00	21:00	01:00	03:00	09:00	13:00		
dinophyta	25.54	25.12	15.25	18.91	11.91	31.83	21.43	7.40
cyanoabcteria	--	--	--	--	--	--	--	--
haptophyta	6.80	0.00	6.79	1.91	6.10	7.70	4.88	3.14
diatom + ochrophyta	33.95	62.13	40.16	58.34	62.74	57.19	52.42	12.25
other algae	33.71	12.75	37.80	20.84	19.24	3.28	21.27	12.87

Table 7-4 Relative algal composition and its alterations across the geophysical circle analyzed under the microscope in order to validate the findings of the algae detection module.

	daylight		night		daylight		Mean [%]	SD [%]
	17:00	21:00	01:00	03:00	09:00	13:00		
dinophyta	0.00	5.06	0.93	0.77	3.31	0.00	1.68	2.06
cyanoabcteria	--	--	--	--	--	--	--	--
haptophyta	3.53	2.53	9.35	7.69	4.13	4.00	5.21	2.68
diatom + ochrophyta	96.47	89.24	89.72	88.46	92.56	95.20	91.94	3.34
other algae	0.00	3.16	0.00	3.08	0.00	0.80	1.17	1.54

7.2.2. Night Cycle

For the night cycle in the harbor of Arcachon, the ALPACA has been integrated into the submersible housing and the whole system was deployed at 3 m depth under the surface. During the cycle, 10 mL of sample were collected at hourly intervals and subsequently analyzed by the ALPACA. The preliminary evaluation was conducted autonomously by the ALPACA and the results were transmitted to the network controller of the host system. Alterations in the biomass concentration

and the algal population were analyzed afterwards, when the cycle has been completed. To validate the results gained from the ALPACA, alterations in the biomass concentration were compared to the CTD sensor from IDRONAUT. In addition, these alterations were compared within the environmental context as shown in Figure 7-5.
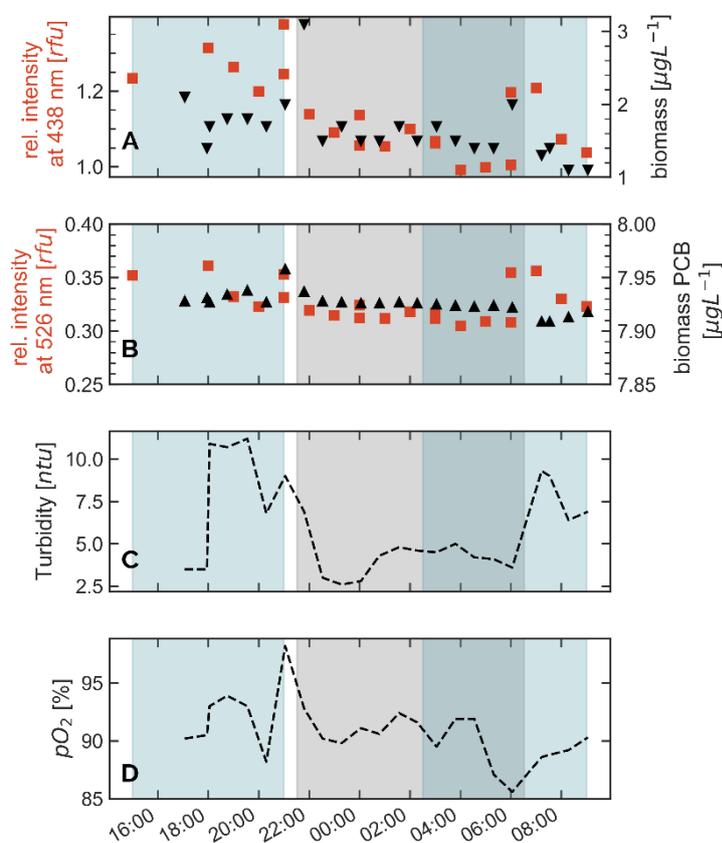
Figure 7-5: Alterations of the total biomass concentration within the environmental context. Alterations in the biomass concentration are recorded by the ALPACA upon excitation at 438 nm (**A**) and upon excitation at 526 nm (**B**) in order to identify algae or cyanobacteria, respectively. In each case, these alterations are verified by the CTD sensor from IDRONAUT (marked in black). In addition, these alterations are compared to further environmental parameters, in particular turbidity (**C**) and oxygen (**D**). Hours of darkness are marked in grey, whereas low tides are marked in light blue.

The ALPACA observes a decrease in biomass concentration during the night, which is verified by the Seapoint chlorophyll fluorometer. At sunset the chlorophyll fluorometer identifies a sudden increase of phytoplankton biomass, which can also be seen on the ALPACA. The ALPACA recognized another increase in the biomass concentration at 6 o'clock during sunrise, however to a smaller extend than before. This increase in total biomass can be verified by the reference as shown in Figure 7-5A, whereas the dynamics for the cyanobacteria (compare Figure 7-5B) cannot be completely confirmed by the reference. There are small outliers during sunrise and at the beginning of the cycle, however the concentration is low in general. Since the turbidity measurement also recognizes an increase during that time, these observations suggest that the ALPACA might suffer from scattering problems. Besides, some alterations might also stem from fluctuations caused by tidal currents.

The oxygen saturation is reduced during night as the phytoplankton consume oxygen during respiration, whereas oxygen evolves during photosynthesis that is performed during the day.

In addition to the semi-quantification of the phytoplankton, the relative algal composition was examined using the LDA. In this case, no validation has been carried out by the University of Genoa. The results are visualized in Figure 7-6 and in Table 7-5. Similar to the previous experiment, diatoms and ochrophytes are the dominant group in the harbor of Arcachon, followed by dinophytes. Cyanobacteria were noticed during the night cycle at midnight and during sunrise.
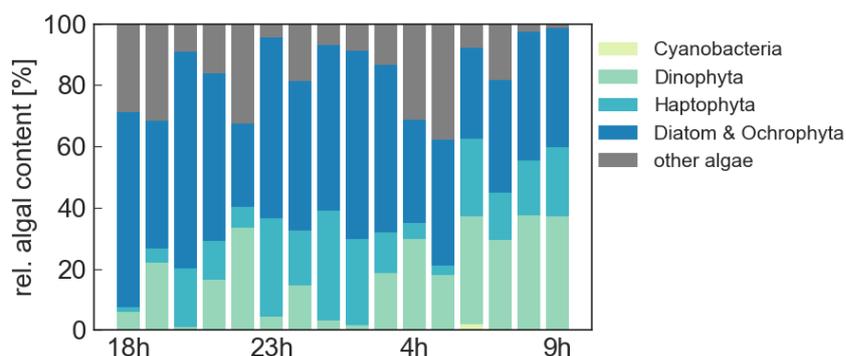
Figure 7-6: Changes in the relative algal composition measured by the algae detection module and examined using the linear discriminant analysis (LDA) as pattern recognition algorithm.

Table 7-5 Alterations of relative algal composition during overnight experiment at Arcachon harbor measured by the submersible algae detection module.

	18:00	19:00	20:00	21:00	22:00	23:00	24:00	1:00	2:00	3:00
dinophyta	5.94	22.09	0.96	16.53	33.53	4.41	14.11	3.22	1.67	18.57
cyanoabcteria	--	--	--	--	--	--	0.52	--	--	--
haptophyta	1.55	4.49	19.1	12.58	6.75	32.26	17.92	35.73	28.23	13.37
diatom + ochrophyta	63.75	41.92	70.96	54.56	27.14	58.86	48.81	54.18	61.32	54.54
other algae	28.76	31.5	8.99	16.33	32.58	4.47	18.64	6.88	8.77	13.53

	4:00	5:00	6:00	7:00	8:00	9:00	Mean [%]	SD [%]
dinophyta	29.69	17.94	35.31	29.45	37.35	37.15	19.25	13.33
cyanoabcteria	--	--	1.99	--	--	--	0.16	0.51
haptophyta	5.28	3.27	25.21	15.57	18.17	22.66	16.38	10.55
diatom + ochrophyta	33.57	41.05	29.72	36.65	41.99	38.71	47.36	12.92
other algae	31.46	37.74	7.75	18.33	2.5	1.48	16.86	12.04

7.2.3. Pollution Gradient in Front of Gujan-Mestras

Gujan-Mestras is the industrial zone of Arcachon, located on the southern waterside of the Arcachon Bay. Here, the environmental conditions are harsh, since the water suffers from industrial waste discharged into the water body and algal populations might differ in the absolute amount and

their relative composition. Therefore, three samples were taken, starting at Pointe de la Matelle (#G1) in front of the harbor, via la Hume (#G2) and finally, the last sampling site was at Chenal de Gujan. Here, two samples were taken within 3 hours (#G3 and #G3F). For further information of the sampling site, please refer to Figure 7-2. All samples were taken at 3 m depth under the surface. This experiment was conducted by the standalone ALPACA included into the submersible probe. Regarding the variations in the biomass concentration, these results were controlled by the TriLux algae sensor from Chelsea Technologies Ltd. The results of the average biomass concentration are shown in Figure 7-7. Both systems indicate that the algal population increases in front of the industrial zone, with low fluctuations over time.
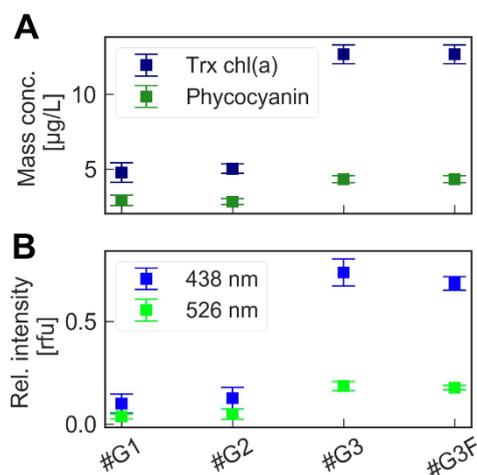
Figure 7-7: Relative biomass concentration recorded either on the TriLux algae sensor from Chelsea Technologies (**A**) or the ALPACA system (**B**). Both systems target the detection of algae and cyanobacteria and both identified an increase of phytoplankton biomass in front of the industrial zone suffering from harsher environmental conditions. While the biomass concentration at Pointe de la Matelle (#G1) was small, the biomass subsequently increases at la Hume (#G2) and was maximal in front of the industrial zone at Chenal de Gujan (#G3 and #G3F).

In addition, the average algal population was evaluated by the LDA as it was previously done. The results are shown in Table 7-6. The overall algal composition is similar to the previous experiments, although there are variations in the relative composition at different sampling sites. In all cases, diatoms and ochrophytes dominate the sample composition, and their relative amounts increase to more than 50% in front of the industrial zone. Dinophytes increase slightly at la Hume, while they decrease under harsh environmental conditions in front of the industrial zone. Similar was the trend for haptophytes, whereas cyanobacteria were absent at any time of the experiment.

Table 7-6 Relative algal content for different sampling sites examined by the standalone ALPACA incorporated into the submersible housing. [a]

	G1	G2	G3	G3F	Mean [%]	SD [%]
dinophyta	15.39	26.34	14.90	14.42	17.76	5.73
cyanoabcteria	--	--	--	--	--	--
haptophyta	7.91	19.36	3.76	0.49	7.88	8.23
diatom + ochrophyta	45.89	38.14	47.80	56.84	47.17	7.68
other algae	30.81	16.16	33.54	28.25	27.19	7.66

[a] Abbreviation: **G1** – Pointe de la Matelle, **G2** – la Hume, **G3** – Chenal de Gujan and **G3F** – Chenal de Gujan after 3 hours.

7.3. Gironde Estuary

The final field trip took place in the Gironde estuary from June 20th to June 25th, 2017 and was similarly structured as the previous field trip in Arcachon. A diurnal cycle was conducted in front of the forest of Royan and samples collected at different sampling sites were analyzed and compared to each other. An overview of the different sampling sites is provided in the map of Figure 7-8. In the estuary, the water body is dominated by seawater organisms than further inside, where it is dominated by freshwater organisms. Since the spectral properties of marine phytoplankton differ from freshwater phytoplankton and since the ALPACA is dedicated to monitor coastal environments, the database of the ALPACA only contains phytoplankton common in marine environments. Thus, it was interesting to examine, how the ALPACA handles unknown phytoplankton samples. Moreover, it is known, that dinoflagellates are present in this estuary during spring and summer, although their relative content might not exceed the content of diatoms and ochrophytes.[168]

Figure 7-8: Different sampling sites for the field trip in Gironde from June 20th to June 25th, 2017. First a geophysical cycle of 24-hours was conducted in front of the forests of Royan (marked in red). Subsequently, three different samples widely spread over the estuary, marked in green, were collected.

A TriLux algae sensor from Chelsea Technologies was used to validate the semi-quantification of biomass concentrations as it was done in previous field trips. However, the University of Genoa did not participate in this field trip, thus a verification of the qualitative analysis was not possible.

7.3.1. Geophysical Cycle

For the geophysical cycle, the forests close to Royan has been selected as appropriate sampling site as here, the water body is dominated by sweater phytoplankton. The ALPACA was enclosed into the submersible probe and deployed on-site at 3-meters depth under the surface. At hourly intervals, 10 mL of cleaning solution and 10 mL of sample were recorded in the measurement chamber and preliminary results, whether algae or cyanobacteria are dominating the samples, were transmitted

to the network controller of the host system. Using the submersible probe, only a single determination was performed. For blank measurement, 70% ethanol was used as common disinfectant to prevent biofouling in the measurement channel. Repetitive controls of the blank level were suitable to maintain the system performance. An overview of the blank measurements is given in Figure 7-9 and in Table 7-7. At the end of the measurement, the blank level increased as the cleaning solution run out and only air was pumped through the system, which must be taken into account for baseline correction during the evaluation. Furthermore, the blank measurements after 6 a.m. were excluded for the averaging process. In these cases, the average blank value of previous measurements was used. Apart from this and an outlier at 21 o'clock, the blank level remains constant during the geophysical cycle.
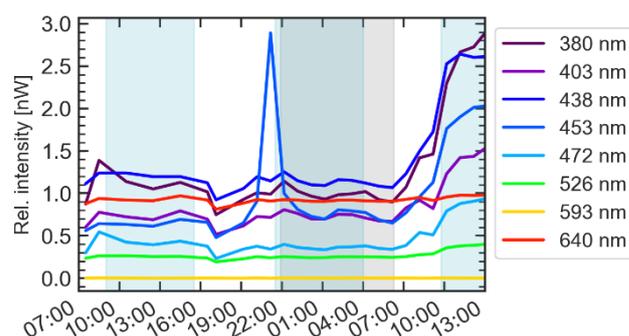
Figure 7-9: Average blank level during the geophysical cycle conduced in the estuary of Gironde in front of Royan. At the end of the cycle, the cleaning solution run out and only air was pumped through the system. Apart from this and an outlier at 21 o'clock due to air bubbles in the system, the blank level remains constant over time.

Table 7-7 Average blank level and its standard deviation during the geophysical cycle for each LED. Due to air in the system, the measurements after 6 a.m. were excluded from the averaging process. Values are given in picowatts.

	380 nm	403 nm	438 nm	453 nm	472 nm	526 nm	593 nm	640 nm
mean	1297.1	830.0	1401.7	1009.4	460.5	269.6	0.2	922.3
deviation	617.7	267.6	540.0	596.6	195.5	52.0	0.9	35.2

Results of the semi-quantification during the diurnal cycle are provided in Figure 7-10. The relative fluorescence intensity of the ALPACA used for semi-quantification was verified by the TriLux algae sensor of Chelsea Technologies. The results are shown in Figure 7-10. The overall trends for chlorophyll-a and phycocyanin plus phycoerythrin concentration recorded on the ALPACA agree well with the reference. Fluctuations due to tidal currents are observed on both systems, on the ALPACA and on the reference sensor. The sudden increase during sunset that was observed in the blank measurement for LED-453, resurfaced during the measurement but was verified by the reference. Compared to the previous field trip one month before in Arcachon, the average biomass concentration has slightly been increased from 5 µg L⁻¹ to 8 µg L⁻¹ (with regard to the chlorophyll-a content in the surface water). In sum, the concentration of chlorophyll-a exceed the concentration of cyanobacteria and rhodophytes, observed upon excitation at 526 nm.
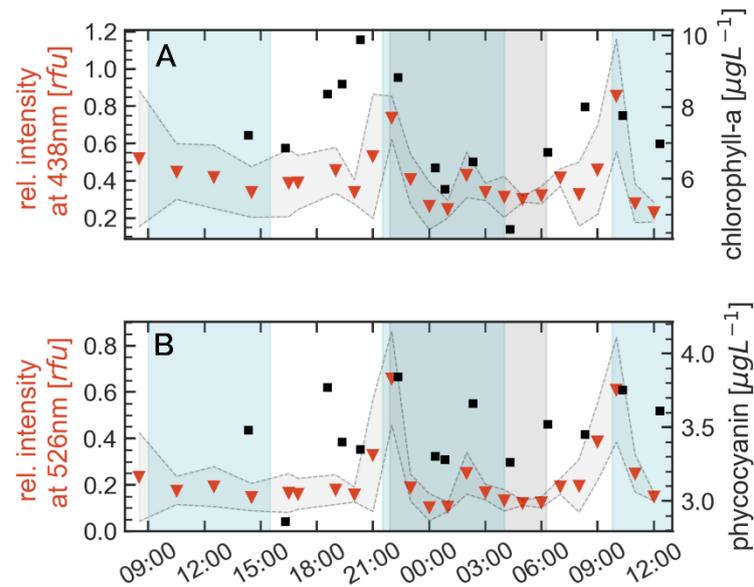
Figure 7-10: Alterations in the biomass concentration recorded on the ALPACA and preliminary results transmitted to the network controller of the host system. The relative trend of the biomass concentration upon excitation at 438 nm targeted the excitation of chlorophyll-*a* and the detection of algae (**A**), whereas the excitation at 526 nm targeted the excitation of phycoerythrin and phycocyanin in order to detect cyanobacteria and rhodophytes (**B**). Alterations in the relative fluorescence intensity during tidal currents are confirmed by the TriLux algae sensor (marked in black), although there are some deviations in the average signal intensity for the LED targeting the identification of cyanobacteria and rhodophytes. Hours of darkness are marked in grey, whereas hours of low tides are marked in light blue.

Table 7-8 Relative algal composition and its alteration over one geophysical cycle at the Gironde estuary. Samples were analyzed at hourly intervals by the APACA incorporated into the submersible probe and deployed on-site. Hours of darkness are marked in bold letters.

	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00
dinophyta	38.20	32.83	24.35	32.12	32.03	34.18	22.12	19.31	43.18
cyanobacteria	--	3.18	10.54	--	--	--	0.48	--	--
haptophyta	25.95	20.46	8.13	26.68	24.75	28.31	10.70	15.76	12.51
diatom + ochrophyta	35.84	30.26	25.83	33.79	39.46	33.51	49.16	36.06	44.31
other algae	0.00	13.27	31.16	7.42	3.75	4.00	17.53	28.87	0.00

	17:00	18:00	19:00	20:00	21:00	**22:00**	**23:00**	**00:00**	**01:00**
dinophyta	42.81	11.20	29.27	27.83	26.46	30.34	36.21	33.97	44.76
cyanobacteria	--	--	--	5.09	--	--	--	--	--
haptophyta	17.89	6.00	23.76	25.07	15.58	31.59	44.15	43.43	29.55
diatom + ochrophyta	39.30	53.84	37.06	35.49	47.18	10.83	17.52	22.18	25.70
other algae	0.00	28.97	9.90	6.52	10.79	27.24	2.13	0.42	0.00

	02:00	**03:00**	**04:00**	**05:00**	**06:00**	07:00	Mean [%]	SD [%]
dinophyta	9.85	34.45	42.77	42.77	31.44	40.89	31.81	9.52
cyanobacteria	--	--	--	--	--	--	0.80	2.39
haptophyta	17.15	33.47	4.19	4.19	1.70	35.77	21.11	12.04
diatom + ochrophyta	41.46	31.48	53.04	53.04	65.76	19.45	36.73	13.03
other algae	31.55	0.60	0.00	0.00	1.09	3.89	9.55	11.50

In addition, after preliminary evaluation, a detailed evaluation using the LDA has been performed subsequently in the laboratory. Results are presented in Table 7-8. In sum, diatoms and ochrophytes are dominating the algal population, followed by dinophytes which are the second largest group in this population. Having a closer look on the algal composition during day and night, diatoms and ochrophytes are dominant in the surface water during the day, whereas dinoflagellates are more prominent during the night. Also haptophytes increase during the night, while cyanobacteria were found only during the morning and before sunset at 8 a.m.

7.3.2. Different Sampling Sites

Finally, the average biomass concentration and the relative algal composition from different sampling sites were compared to each other. In particular, differences between Royan, located at the inner site of the estuary, and the open waters were in the focus of attention. The sampling sites are shown in Figure 7-8. Again, the ALPACA was incorporated into the submersible housing and deployed at 3-meters depth under the surface.
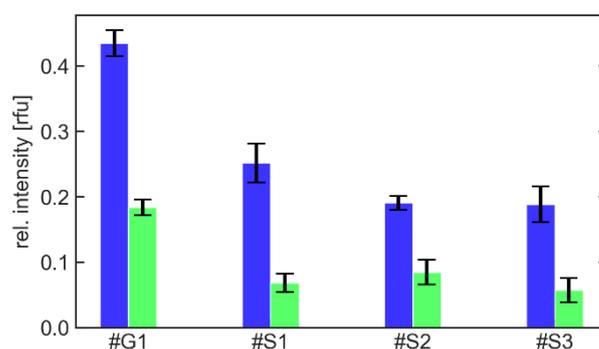
Figure 7-11: Relative biomass concentration at different sampling sites widely spread over the Gironde estuary. In the first sampling site (#G1) the biomass concentration is determined during the geophysical cycle in front of the forests close to Royan at the same time of the day, whereas the other sampling sites are closer to the open waters.

As assumed and as demonstrated in Figure 7-11 and Table 7-9 the biomass concentration decreased close to the open waters compared to the estuary. The relative algal composition, evaluated by the LDA, varies at all different sampling sites, although the relative content of dinophytes is high in each case. At sampling sites close to the estuary, dinophytes even dominate the algal population with 36.2–40.3%. Although, diatoms plus ochrophytes are most common in this areas, their relative content only slightly exceeds the relative content of dinophytes and haptophytes in the estuary (#G1) and in the open waters (#S3). Cyanobacteria were found to a minor extend in the estuary, whereas their relative content increased in front of the ocean (#S3) to even more than 10%. Unfortunately, these results weren't validated by the microscopic analysis since the University of Genoa didn't participate in this field trip.

Table 7-9 Relative algal composition at different sampling sites widely spread over the Gironde estuary. The composition was evaluated by the LDA and the values are given in percent. The first sampling site (#G1) is the average level of the geophysical cycle in front of the forests of Royan, whereas the other sampling sites are farther outside in the open water.

	#G1	#S1	#S2	#S3
dinophyta	31.81	40.29	36.21	24.35
cyanoabcteria	0.80	--	--	10.54
haptophyta	21.11	26.21	35.69	8.13
diatom + ochrophyta	36.73	29.90	28.10	25.83
other algae	9.55	3.60	--	31.16

Part III

Results targeting the Hybrid Sensing

8. Graphical user Interface dualSense for the Evaluation of Hybrid Sensors

For simulation and evaluation of the hybrid sensors, a software algorithm, called dualSense, was written in Python, which bases on the analysis of two different analytes by the measurement of their overall phase shifts at two different modulation frequencies. In this chapter, the operating instructions for the software are presented, whereas the different mathematical principles for evaluation of the single-layer dual sensor and the double-layer dual sensors, including their error estimation, are presented in section 3.3.

8.1. Software Installation and Directory Structure

The dualSense software is written in Python and an executable version of the dualSense Software is provided in the Appendix 11. The graphical interface of the software opens after double-click on the corresponding software icon.

It is advised to maintain the structure of the subfolders and the nomenclature of the calibration files as it is initially given. On account of clarity, the directory structure is kept forthrightly and consists of two subfolders – the '*calibration*' folder and the '*reports*' folder. The '*calibration*' folder contains all required calibration files and must be kept up to date for the specific sensor. Prior to modifications of the calibration files, the user ought to get used to the structure and the information of the calibration files. In particular, the calibration file for the pH/T dual sensor contains a temperature calibration of all Boltzmann parameters, whereas the calibration files for the pH/O2 or the CO2/O2 dual sensors contain the calculation factors for the conversion between the phosphorescence lifetime and the relative phosphorescence intensity along the oxygen concentration. In contrast, the calibration file for the single-layer O2/T dual sensor is structured in a more complex manner as it contains all calibration measurements of the specific sensor. Besides the specific analyte content at certain calibration points, the intensity ratio, the phosphorescence lifetime and their corresponding normalizations are given at the calibration points. In contrast, the '*reports*' folder contains further subfolders for the error propagation and the processed results of the dual sensor simulations or evaluations.
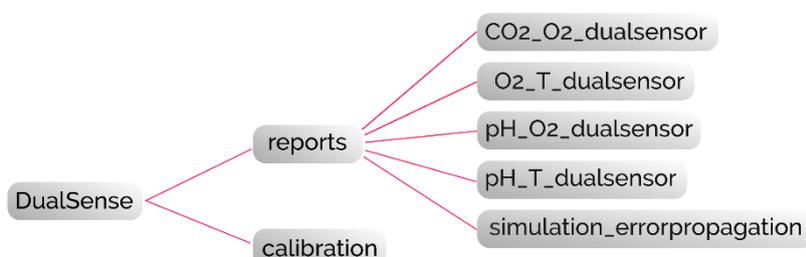
Figure 8-1: Recommended directory structure for the dualSense software interface

8.2. Software Interface and general Operating Information

After installation and starting of the software, the graphical user interface will appear on the screen as it is shown in Figure 8-2.
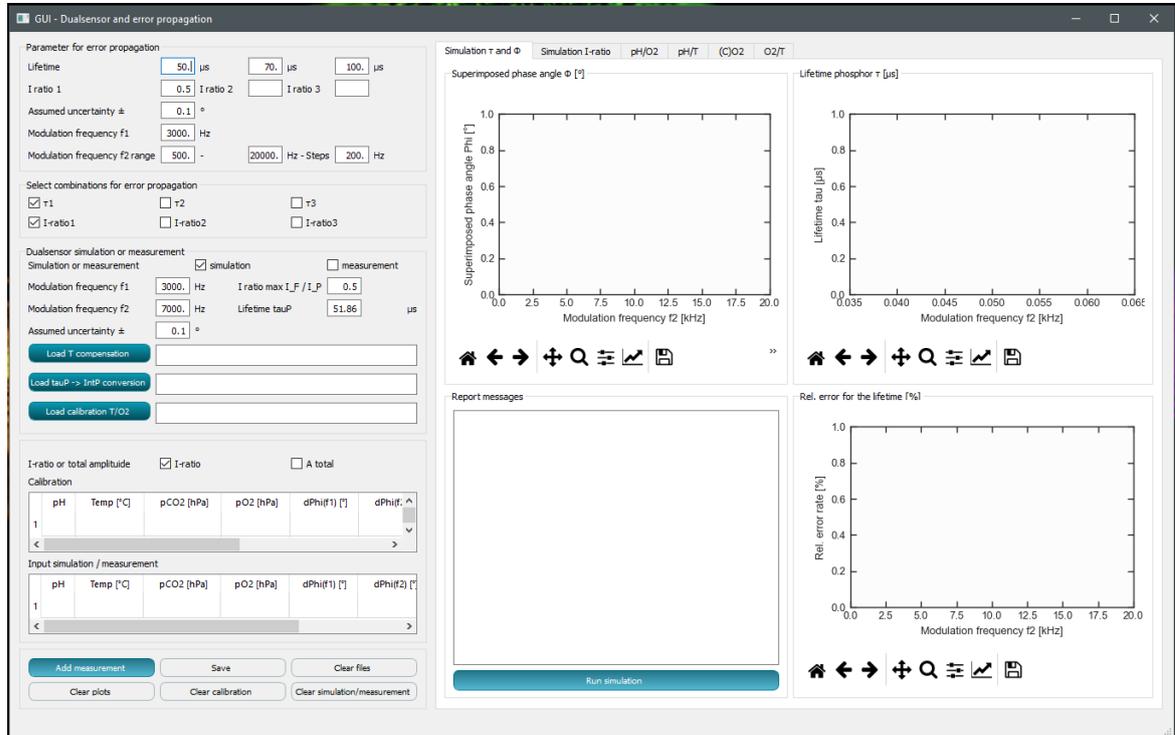
Figure 8-2: Graphical user interface for the dualSense software used for the estimation of errors that occur during simulation of evaluation. The left panel is dedicated to the operational parameters and calibration or compensation files, while the right panel is dedicated to the presentation of the computed results. The results are shown in the corresponding plots and in the message box in a summary form.

In general, the software interface is divided into two parts. On the left side are all operating parameters required for investigation of the error propagation and the simulation or evaluation of the dual sensors, whereas the right side is intended for the presentation of the results and their visualization.

The left side on the top is dedicated to investigations of the error propagation that occur during simultaneous analysis of two different analytes using a modified dual-frequency analysis (see Figure 8-3). The user must define the operating parameters such as the modulation frequencies and the assumed measurement uncertainty of the phase angle. The first modulation frequency is fixed, whereas an operating range is defined for the second modulation frequency, within which the frequency is optimized in order to reduce the overall error estimation. Besides the operating parameters, the intrinsic sensor properties, such as the lifetime of the long-lived dye and the intensity ratio of the fluorescent and the phosphorescent dye, must be defined. For the purpose of comparison, multiple decay times and intensity ratios can be defined.

Below these parameters, the user can additionally define the combinations of decay time and intensity ratio, that should be computed with a maximum number of four different combinations.

Figure 8-3: Operating parameters for investigations of the error estimation and the optimization of the second modulation frequency.

The cells below are dedicated to the simulation and evaluation of the dual sensors using the two frequency analysis (see Figure 8-4) First, the user defines whether the phase shifts of the dual sensor should be simulated for one theoretical measurement point or whether the phase shifts of the dual sensor should be analyzed at one or multiple measurement points. Further, the operating parameters, such as the modulation frequencies, the maximal intensity ratio of the dual sensor and the average measurement uncertainty of the phase angles, have to be defined by the user.

Below, further txt-files required for the individual sensor calibration and compensation can be loaded. In particular, the dual sensor pH/T requires a temperature compensation that can be loaded by clicking the '*Load T compensation*' button. The dual sensors pH/O2 and CO2/O2 however, require conversion factors that enables the translation between the phosphorescent decay time into the signal intensity at any oxygen concentration. Since the analysis of the single-layer sensor T/O2 is built on a different principle than the dual-layer sensors, a more complex calibration file is required for this hybrid sensor. The calibration and compensation files can be erased by clicking the '*Clear file*' button on the bottom.

In the following tables, the content of the corresponding analytes at different calibration or measurement points has to be defined. For the simulation / measurement table, only one row is initially shown, as only one theoretical point can be computed during simulation. For the evaluation of the measurement and since the software should provide monitoring capabilities, multiple measurement points can be analyzed and evaluated. Further rows can be added by clicking the '*Add measurement*' button. In addition, the user can decide whether the overall amplitude signal at two different modulation frequencies was measured or whether the specific intensity ratio was available. The individual tables can be erased by clicking one of the clearing buttons at the bottom of this panel, either the '*Clear calibration*' button or the '*Clear simulation/measurement*' button.

Figure 8-4: Widget for the simulation or evaluation of different dual sensors. Besides the operating parameters, the user must define the calibration points and have to refer to the corresponding calibration and conversion files for a specific dual sensor. In addition, one theoretical measurement point for the simulation had to be defined. For the evaluation of a realized measurement, the user can enter data for multiple measurement points.

If the presented results on the right panel (plots and message box) should be erased for any reason, press *'Clear plots'*. By pressing this button, only the figures and message box in the open tab will be cleared. Other plots and summarizing boxes will remain.

After simulation or analysis of the measurement, the results can be stored within a synoptic txt file by pressing the '*Save'* button. After pressing '*Save'*, a '*reports'* folder and corresponding subfolders for the specific project are generated next to the graphical user interface (GUI) in the same directory, if the subfolders do not yet exist. Please note, that only the input parameters and the results will be saved within this txt file. If requested, the calibration and evaluation plots have to be saved separately.

In the following subsections, the individual tabs dedicated to the error propagation and different dual optical sensors, are explained in more detail. However, for more information on the mathematical procedures, reference is made to chapter 3.

8.2.1. Error Propagation

The first and the second tab are dedicated to the simulation of the overall phase shifts of a dual optical sensor system and the estimation of the maximal error due to a linear error propagation across the evaluation procedure presented earlier in section3.3. On the right panel, the user can define up to three different decay times and the maximal intensity ratio of a theoretical dual optical sensor. In the section below, the combinations of decay time and intensity ratio, that ought to be computed, can be defined by the user. For the purposes of clarity, up to two different decay times and intensity ratios can be selected at the same time, resulting in four combinations and eight different overall phase shifts.

Besides, the expected measurement uncertainty in the phase shift has to be defined as well as the modulation frequencies that are required for the optimization procedure. Initially, the first modulation frequency is fixed since one modulation frequency is mostly known. However, the operating range for the second modulation frequency (start, stop and slit width) has to be defined with respect to the technical capabilities.

By pressing 'Run $simulation$', the overall phase shifts of the theoretical dual sensor and the estimated error propagation are computed for all possible pairs of modulation frequencies. The resulting phase shifts are displayed on the screen as it is exemplarily shown in Figure 8-5. The fixed modulation frequency f_1 produces a constant phase shift $\Phi(f_1)$ (marked in bright blue in Figure 8-5), while the second variable modulation frequency produces a variable overall phase shift $\Phi(f_2)$ (marked in dark blue, see Figure 8-5).
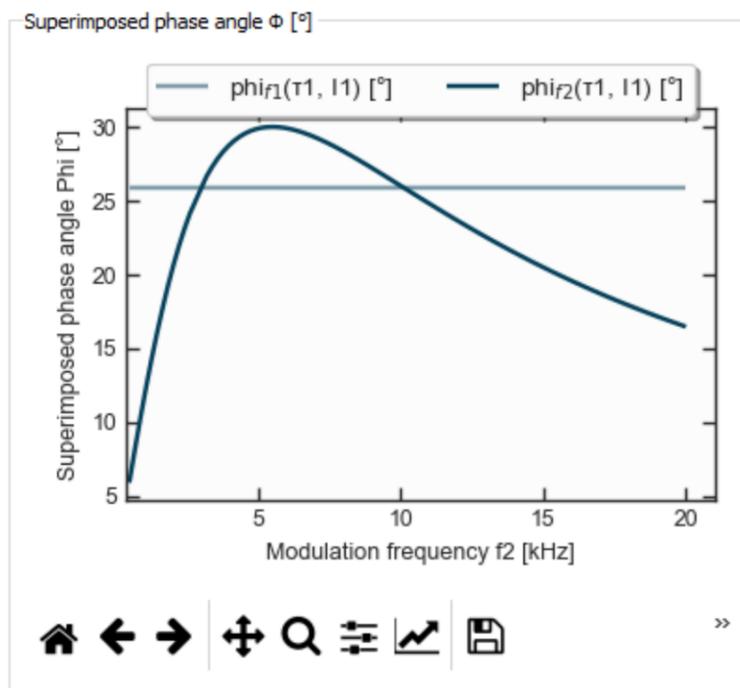
Figure 8-5: Anticipated overall phase shift at different modulation frequencies according to the given operating parameters and the physical properties of the dual sensor. While the first modulation frequency is fixed, the second modulation frequency is scanned between the given operating range.

Besides the overall phase shifts at two modulation frequencies, the decay time of the long-lived component and the relative error rate originated from the proclaimed measurement uncertainty in the phase shifts is computed as shown in Figure 8-6. A synoptic overview of the input parameters and relevant temporary and final results is displayed in the message box and can be saved by pressing '*save*'.
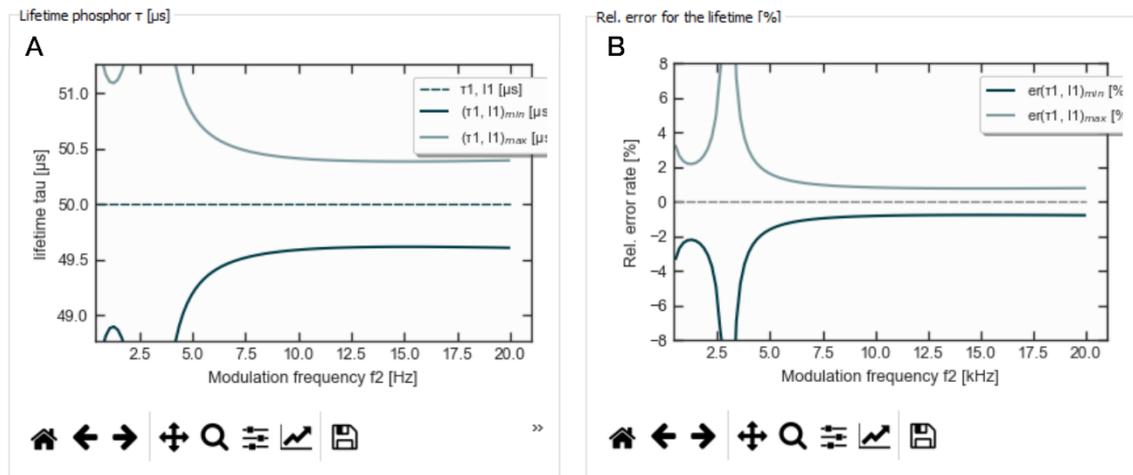
Figure 8-6: Computed decay time of the long-lived dye according to the given operating parameters and the physical properties of the theoretical dual sensor (**A**) and the associated relative error rate (**B**). Both parameters are scanned for different pairs of modulation frequencies in order to minimize the estimated error.

In order to optimize the error propagation and the dynamics of the overall phase shifts, the first modulation frequency can be set to the optimal modulation frequency, at which the calculated phase shift is maximal. The second modulation frequency can subsequently be optimized by pressing the button '*Run simulation*' again. Thus, the overall process of minimizing the error propagation for an ideal pair of modulation frequencies is an iterative process.

A corresponding procedure is additionally conducted for the intensity ratio of the long-lived dye compared to the fluorescence dye on the second tab. Here, the error propagation is investigated for the intensity ratio and the optimal pair of modulation frequencies is determined where the error propagation is minimal. In general, the intensity curve follows the same trend as the decay time and in sum, a modulation frequency that is at least twice of the first modulation frequency might be preferred. Here, the maximum of 20 kHz is originates from technical limitations of the measurement setup, however they can be defined in the left panel (see Figure 8-3).

8.2.2. Dual Optical Sensor for pH and Oxygen

The next tab of the software interface is dedicated to the two-layer sensing system targeting the simultaneous analysis of pH and oxygen. This dual optical sensor requires two calibration points for each analyte and an additional conversion file, which enables the conversion of the decay time

of the long-lived dye into the corresponding phosphorescence intensity at any oxygen concentration. The conversion file is generated after careful calibration of the individual optical sensor layer responsible for the oxygen sensing. It is assumed that the individual sensor remains unchanged without any reciprocal interference, when the individual sensor layers are combined to a dual optical sensor. In addition, several fitting parameters and their default values are already set, in particular the modulation frequencies, the assumed measurement uncertainty for the phase shifts as well as the maximal intensity ratio and the maximal decay time of the long-lived dye. Moreover, the default values for the point of inflection (pK_a also referred to as V50) and the slope for the pH sensing and the K_{sv}-value, the proportionality coefficient and the slope for the oxygen sensing are already set, although these fitting parameters can be adjusted in the lower part of the right panel, if required. An example of the software interface is shown in Figure 8-7. In this example,
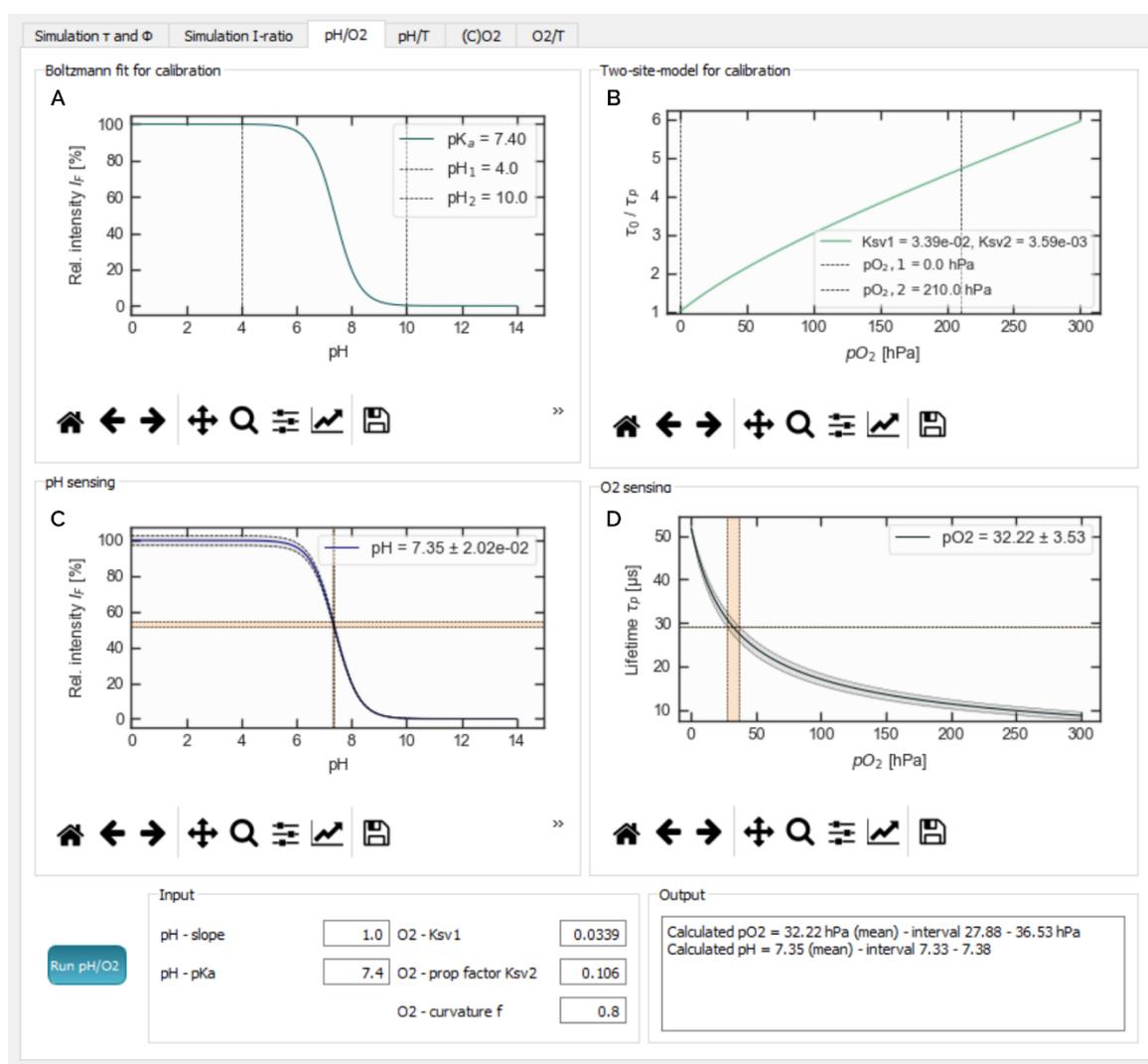
Figure 8-7: Software interface dedicated to the dual optical sensor aiming at the simultaneous analysis of pH and the oxygen concentration. The upper part is intended for calibration purposes: (**A**) Based on the calibration points for the pH sensing, the Boltzmann sigmoid is calculated with respect to the given fitting parameters. (**B**) For the oxygen sensing, the two-site-model is used for calibration and fitting purposes. Relevant fitting parameters are defined below in the input section and can be adjusted, if required. (**C-D**) The lower part of this tab is destined for the simulated measurement point or for measurement evaluation, including the determined error propagation. Below, the calculated results are represented within the message box.

the dual optical sensor has computed a theoretical measurement point and its estimated error propagation.

If a theoretical measurement point should be simulated, the desired pH value and the oxygen concentration have to be defined in the input table on the left. By pressing the '*run*' button, the individual calibrations for the pH sensor and for the oxygen sensor are computed with respect to the Boltzmann equation and the two-site-model, as described in section 3.3.4. The phase shifts and the intensity ratio for all calibration and theoretical measurement points at different modulation frequencies are inserted into the corresponding tables. The estimated error, originating from the measurement uncertainty in the phase shifts, is considered during simulation and reported in the message box.

While only one theoretical point can be calculated during simulation, multiple measurement points can be analyzed during evaluation. In order to evaluate several measurement points, add further rows to the input table. This can be realized by clicking the '*Add measurement*' button. For the subsequent measurement evaluation, the overall phase shifts and the overall signal amplitudes at two different modulation frequencies are required in addition to the general operating parameters. After pressing the '*run*' button, the phase shifts and the overall amplitudes are analyzed and the results are displayed on the screen. If desired, the results can be saved in a txt file, while the plots have to be saved separately.

8.2.3. Dual Optical Sensor for pH and Temperature

The subsequent tab of the software interface targets the simultaneous analysis of the pH and the temperature as dual-layer sensor system. Similar to the previous dual sensor, the pH sensitive dye is calibrated according to the Boltzmann fit and requires two different calibration points. In addition, also the temperature sensitive layer requires two different calibration points. According to the first sensors, the temperature sensing is restricted to a temperature range of 0 - 50°C. For temperature compensation of the pH sensor, the individual pH sensor layer and, in particular, the fitting parameters must be calibrated along the temperature range in advanced. It is assumed, that the individual sensing layers remain unchanged without reciprocal interferences, when the individual sensing layers are combined to the dual sensor, which is why the determined fitting parameters of the individual sensor layer can be adopted to the dual optical sensor. In each case, the user can decide for which fitting parameter the temperature compensation ought to be considered. During investigation, it was noticed that the slope of the Boltzmann sigmoid is independent of the temperature and can therefore be neglected during fitting and its average value is used for calibration. In addition to the calibration, the modulation frequencies have to be defined. The initially tested dual sensors require a modulation frequency of 3000 Hz, whereas the second modulation frequency can be optimized upon request. An example of the software interface dedicated to the simultaneous determination of the pH and the temperature is shown in Figure 8-8.

Herein, the dual optical sensor has computed a theoretical measurement point and its estimated error propagation. Apart from the slope, all remaining parameters of the Boltzmann sigmoid are temperature compensated.

For simulation of a theoretical measurement point, the desired pH value and the temperature ought to be defined in the input table on the left. After pressing the '*run*' button, the pH sensor is calibrated according to the Boltzmann sigmoid, while the selected parameters are additionally temperature compensated. For the temperature sensor a linear regression is performed. Further information on the mathematical procedures is provided in section 3.3.4. The extrapolated phase shifts for all calibration and simulation points at both modulation frequencies are inserted into the corresponding tables. The determined standard deviation, originated from the measurement uncertainty in the phase shifts and due to the error propagation during simulation, is displayed in the message box.
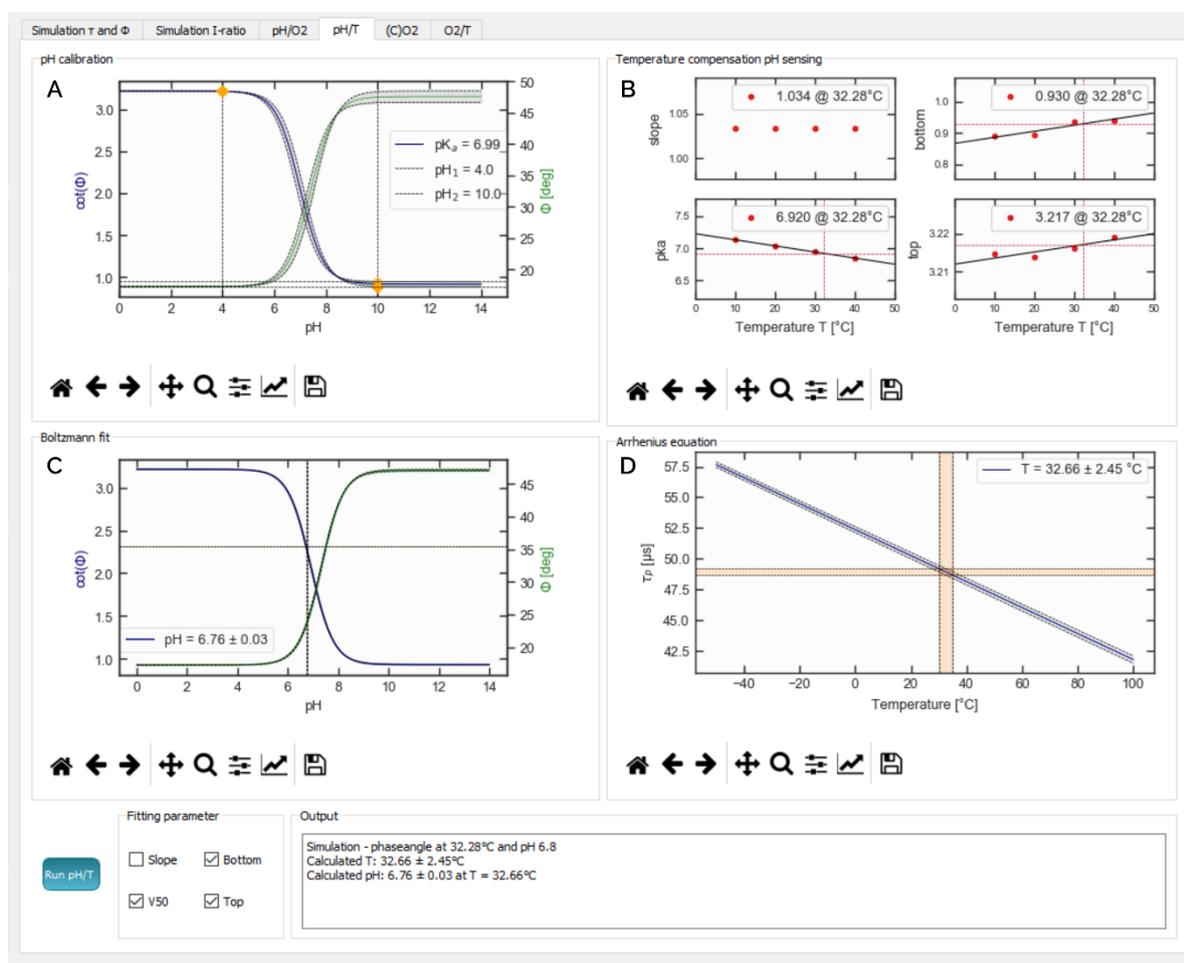
Figure 8-8: Software interface dedicated to the simultaneous sensing of pH and temperature. The upper part is intended for calibration purposes: (**A**) Based on the calibration points for the pH sensing, the Boltzmann sigmoid is calculated with respect to the given fitting parameters. (**B**) For the temperature compensation of the pH sensor, a corresponding calibration file is required that contains the temperature calibration of the fitting parameters for the Boltzmann sigmoid. The user can define in the lower part of this tab, which Boltzmann parameters ought to be temperature compensated. (**C-D**) The lower part of this tab is destined for the simulated measurement point or for measurement evaluation, including the determined error propagation. Moreover, the calculated results are displayed within the message box.

For the analysis of conducted measurements, multiple measurement points can subsequently be evaluated. In order to add further measurement points to the input table, click on the '*Add measurement'* button. For accurate evaluation of the measurement according to the two frequency evaluation, the overall phase shifts are required at two different modulation frequencies. By clicking on the '*run'* button, the given phase shifts are analyzed under consideration of the provided temperature compensation file. The processed measurement results can be saved in a txt file upon request, while the plots have to be saved manually.

8.2.4. Dual Optical Sensor for Carbon Dioxide and Oxygen ©O2

The next tab is dedicated to the dual optical sensor targeting the simultaneous analysis of carbon dioxide and oxygen as dual-layer sensor system. An example of the interface, is shown in Figure 8-9. The sensor is calibrated in ambient air and with an additional, anoxic calibration point containing a definite amount of carbon dioxide. In general, the sensor layers can be calibrated within a carbon dioxide and oxygen range of 0-50 hPa and 0-200 hPa, respectively. Additional operating parameters of the individual sensor dyes are set as default values for regression, however they can be adjusted, if requested. Moreover, a conversion file, enabling the translation between the decay time of the long-lived sensor dye and its phosphorescence intensity, is required. The fitting of both sensor layers is computed according to the two-site-model with respect to the given fitting parameters. The user can decide whether the K_{sv}-values or the decay time of the long-lived dye ought to be optimized during the fitting procedure.

Similar to the previous dual sensors, a theoretical measurement point including their error propagation can be simulated. The overall phase shifts and the amplitude at two different modulation frequencies or the corresponding intensity ratio is calculated and inserted into the tables on the left side and the results are displayed on the screen. However, for the evaluation of (multiple) measurement points, the overall phase shifts and the amplitudes at two different modulation frequencies are required. After ticking the measurement checkbox and pressing the '*run'* button, the measurement points are analyzed and the results are shown in the upper two plots and in the message box as synoptic overview of the measurement results.
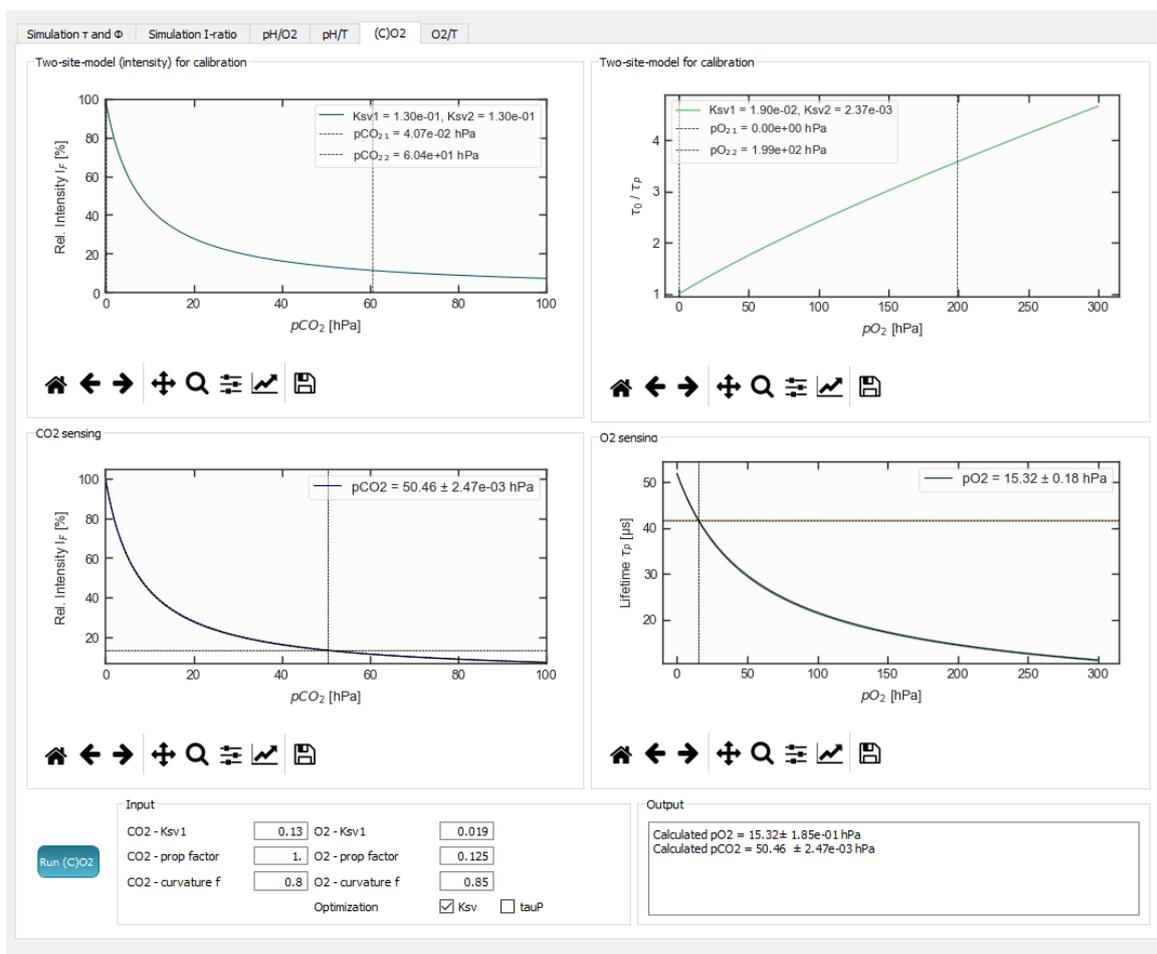
Figure 8-9: Software interface dedicated to the simultaneous sensing of carbon dioxide and oxygen in the gas phase. The upper part is intended for calibration purposes: (**A-B**) Both sensing dyes are fitted following the two-site-model under consideration of the calibration points and the provided fitting parameters. For the oxygen sensing dye, an additional calibration file for conversion between the decay time and the signal intensity of the long-lived component is required. In addition, the user can define whether the K_{sv} value or the decay time is fitted for the oxygen sensing layer. (**C-D**) The lower part of this tab is destined for the simulation of a theoretical measurement point or for the evaluation of a conducted measurement. Besides the visualization of the computed results, the results and the determined error propagation, are reported are reported within the message box.

8.2.5. Dual Optical Sensor for Trace Oxygen and Temperature in a Single-Layer System

This dual sensor comprises a single-layered system, where one sensor is sensitive to two different analytes – in this case, the sensor is sensitive to the oxygen content and the temperature. As the mathematical principle is more complicated than the dual-layer systems, the sensor behavior cannot be simulated similarly to the previous sensors and is rather based on accurate calibration measurements and subsequent extrapolation and regression. Hence, it is not possible to simulate a theoretical measurement point, however one or multiple measurement points can be analyzed with respect to the carefully calibrated dual sensor. As mentioned before, the calibration file has to be loaded using the '*Load calibration T/O2*' button. In addition, the operating parameters have to be defined and the overall phase shifts and the amplitudes at two different modulation frequencies have to be entered into the measurement table. The operating range as well as an appropriate slit width can be defined for both analytes on the bottom of this tab.
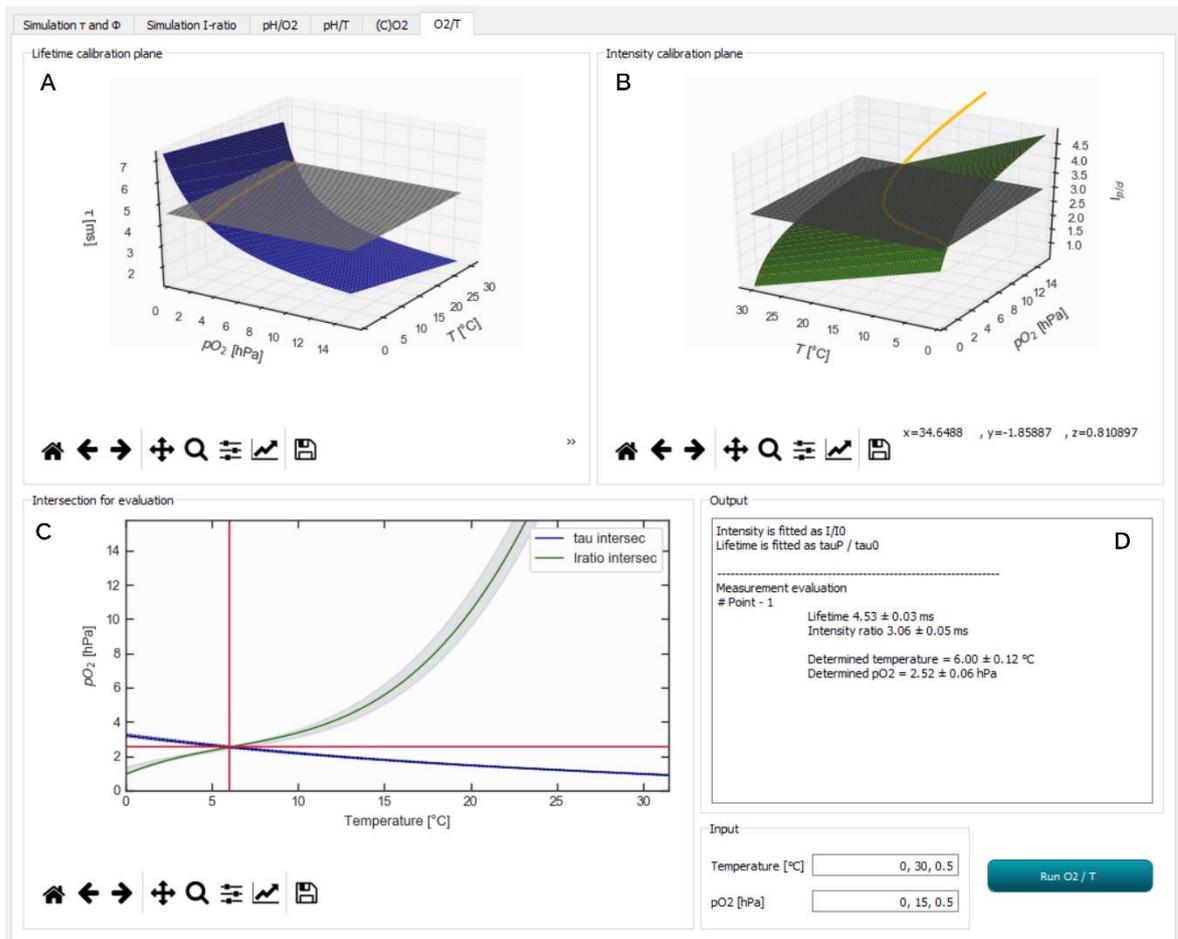
Figure 8-10: Software interface dedicated to the single-layer dual sensor aiming at the simultaneous analysis of the temperature and the oxygen concentration in gas phases or aqueous solutions. (**A**) Based on appropriate calibration measurements, a three-dimensional surface can be modulated for the decay time of the dual sensor at any oxygen concentration and temperature defined within the operating range. (**B**) Moreover, a three-dimensional surface for the intensity ratio of the prompt and delayed fluorescence of the dual sensor is computed. Additionally, the decay time and the intensity ratio of the measurement point is calculated and shown as grey plane in the corresponding plots. The intersection lines are highlighted in orange (decay time) or yellow (intensity ratio). (**C**) Two-dimensional projections of the intersection lines are shown at any oxygen concentration and temperature and their junction is further highlighted in red cross line and a synoptic overview of the calculated results is presented in the message box (**D**).

By pressing the run-button, the calibration measurements are extrapolated and fitted in order to generate a calibration surface for the decay time of the long-lived part of the dual sensor. In addition, a calibration surface for the intensity ratio of the prompt and delayed fluorescence of the dual sensor is computed, as it is demonstrated in the three-dimensional surface plots, shown on top of Figure 8-10A-B. Besides the calibration surface, a grey plane is shown representing the calculated decay time and the intensity ratio of the dual sensor at the measurement point. The intersecting line of the measurement point and the calibration surface is highlighted as orange (decay time) or yellow (intensity ratio) lines. In Figure 8-10C, a two-dimensional projection of the intersection lines and their junction is shown, defining thus, the resulting oxygen concentration and temperature at the measurement point. It has to be highlighted that numerical analysis strongly depends on the quality of the calibration measurements and the defined operating range of the two analytes. Besides the graphical analysis, the intersection point is calculated numerically according to the mathematical descriptions in section 3.3.4. A synoptic overview of the determined decay time, the intensity ratio and the calculated oxygen content and temperature is shown in the message box (see Figure 8-10D).

9. Preliminary Experiments for the Single-Layer Sensor Sensitive to Oxygen and Temperature
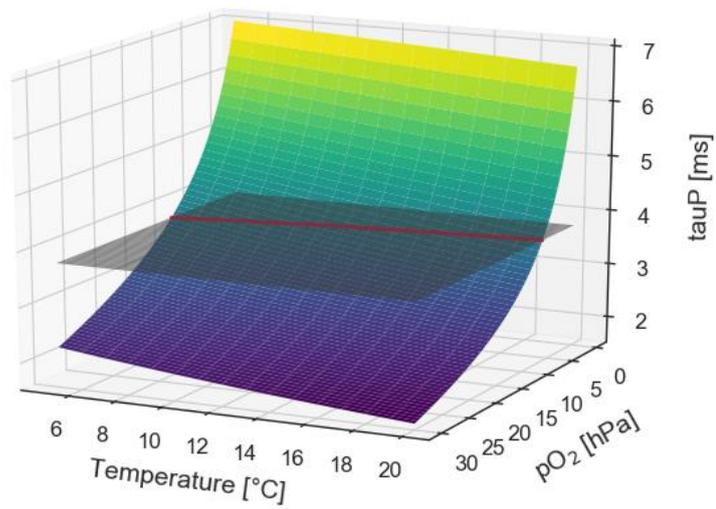
The dual optical sensor enabling the simultaneous sensing of oxygen and temperature in only one sensor layer was realized using a porphyrin dye embedded into a PVC matrix based on the concept of thermally-activated delayed fluorescence (TADF).[104] A publication describing its synthesis, the relevant chemical and physical properties and its evaluation is planned. Thus, only the preliminary measurements conducted in the gas phase and the extrapolation of these measurements in order to generate a reference matrix are described in this chapter.

9.1. Measurement Setup

In order to generate a comprehensive reference matrix of the dual optical sensor, calibration measurements were conducted in the gas phase at different temperatures. For this purpose, the porphyrin-based dual optical sensor was integrated into a glass capillary, which was deployed into a cryostat from Avantor (USA, us.vwr.com) for temperature control. A schematic draft of the measurement setup is shown in Figure 9-1. The in- and out-lets of the capillary were connected to gas pipes using Teflon tubes as connectors and the gas pipes were deployed into the cryostat for preconditioning purposes. The inlet of the gas pipe was further connected to a mass flow controller from msk (USA, www.mksinst.com) to mix the purified nitrogen 5.0 and compressed air (oxygen content 20.95%), according to the preliminary defined gas flows. Both reference gases were obtained from Linde (Austria, www.linde-gas.at). Considering the gas flow and the humidity of the gases, the oxygen concentration was amended between 0 - 100 hPa with steps of 5% for one calibration. The actual partial pressure and the temperature were controlled with a FirestingO2 reference sensor and a high precision digital thermometer, respectively. The reference sensors were obtained from PyroScience (Germany, www.pyroscience.com) and the thermometer from (Germany, www.omega.de), respectively.
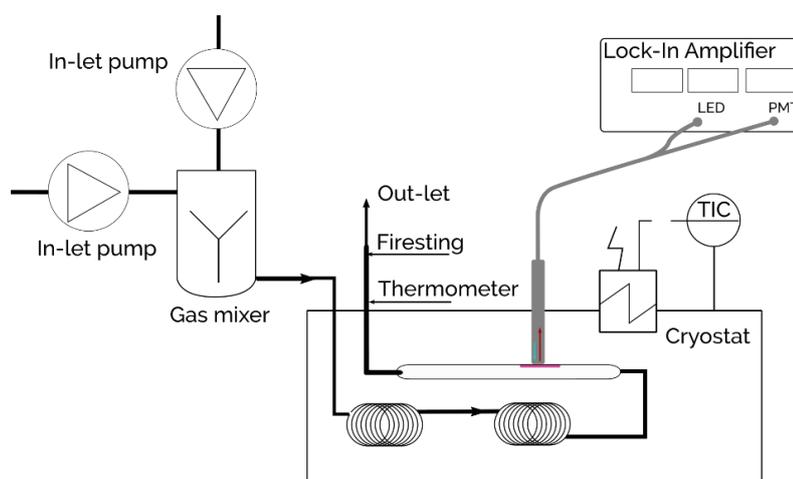
Figure 9-1: Scheme of the measurement setup for simultaneous sensing of oxygen and temperature in the gas phase. The dual sensor is integrated into the glass capillary (shown in pink) and deployed into the cryostat. The in- and out-let of the capillary are connected with the gas pipes which transport the gas from the mixer to the sensor. For signal read-out, the fiber optics is connected to a lock-in amplifier enabling the recording of the overall amplitude and the phase shift.

For read out of the sensor signal, a SR830 DSP Lock-In Amplifier from Stanford Research Systems (USA, www.thinksrs.com) was used at two different modulation frequencies (13.76 Hz and 63.23 Hz, respectively). The initial phase shift of the amplifier was determined for zero adjustment and phase correction at both modulation frequencies using lumogen orange as fluorescence standard. For excitation of the dual sensor, an LED-493 from Roithner Lasertechnik (Austria, www.roithner-laser.com) was combined with two layers of the plastic filter 'Lagoon Blue' obtained from Lee Filters (UK, www.leefilters.com). On the emission side, an RG-645 longpass filter from bk Interferenzoptik (Germany, www.interferenzoptik.de) covered the photomultiplier of the lock-in amplifier. Further operating parameters of the lock-in amplifier were set to the default values. A grounded alternating current was chosen as signal input, the time constant was set to 8 dB and 100 ms and the sensitivity was set to 200 mV. It has to be highlighted that no notch filters must be selected on the lock-in amplifier. Since the modulation frequencies are small with 13.76 Hz and 63.23 Hz, the filters would screen the important signal out. The gas flow of the calibration gases and further operating parameters are defined within the software GLT master control. Within the software, the amplitude of the LED is set to 5 V, whereas the exposure time is set to 2 s with a measuring interval of 5 s. Each gas plateau was kept for 10 min for equilibration of the dual optical sensor.

9.2. Calibration Measurement in the Gaseous Phase

Using the software GLT master control, the gas flow of the reference gases was defined for each plateau according to Table 9-1 with an overall gas flow of 200 mL min^{-1}. For verification, the actual partial pressure was monitored by the FirestingO2 reference sensor. The corrected phase shift and the overall amplitude were monitored over time by the lock-in amplifier. The calibration was successively conducted at two different modulation frequencies (13.76 Hz and 63.23 Hz, one after the other) with a threefold replication at the same temperature before the next temperature was adjusted. Overall three different temperatures were calibrated ranging between 5 - 20 °C.

Table 9-1 Gas flow of the nitrogen and oxygen reference gasses and the resulting oxygen partial pressure (calculated and actual, respectively).

Gas flow N_2 (mL min^{-1})	Gas flow O_2 (mL min^{-1})	Partial pressure pO$_2$ calculated (hPa)	Partial pressure pO$_2$ real (hPa)
100	0	0	0
99	1	1.99	1
95	5	9.94	7.68
90	10	19.89	14.82
80	20	39.77	28.69
70	30	59.66	42.38
60	40	79.54	56.77
50	50	99.43	71.44

9.3. Extrapolation of a Reference Matrix and Evaluation

Afterwards, the average phase shift and amplitude were determined for each oxygen partial pressure. Subsequently, the (relative) decay time and the corresponding intensity ratio of the dual optical sensor were calculated at any oxygen partial pressure and at any temperature according to the mathematical description in chapter 0. An overview of the conducted calibration measurements and the resulted parameters is shown in Figure 9-2A-B.
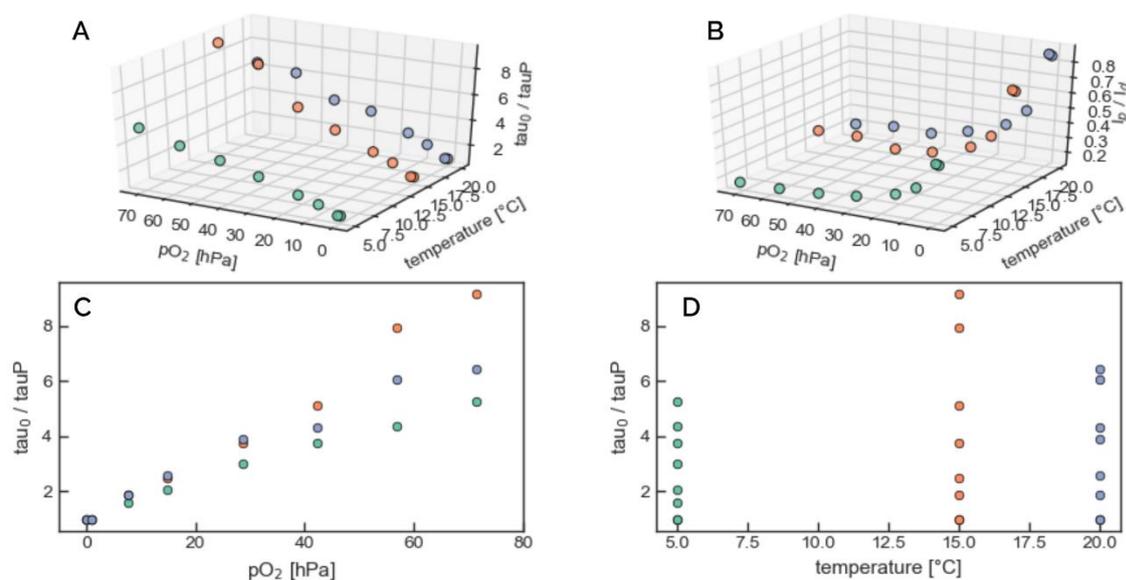
Figure 9-2: Calibration measurements of the dual optical sensor sensitive to oxygen and temperature. Above the calculated relative decay time τ_0/τ_P and the calculated intensity ratio I_d/I_p is displayed against both analytes (**A**-**B**). Below, the decay time is plotted against the actual oxygen concentration (in hPa) (**C**) or respectively against the temperature (**D**) as two-dimensional projection.

An analysis of the mathematical dependency of both parameters – the relative decay time and the intensity ratio in terms of delayed fluorescence compared to prompt fluorescence – was conducted. Thus, it has been evidenced that the behavior of both parameters was consistent within an oxygen concentration of 0-30 hPa, wherefore both parameters were fitted within this concentration range using a two-site-model as regression model. However, in order to describe the temperature dependency, a linear regression was estimated for the fitting parameters within the two-site-model. The temperature calibration was considered within a temperature range of 5-20°C. The final reference matrix for the dual optical sensor analyzed in the gas phase is shown Figure 9-3.

For evaluation of a particular measurement point, the actual decay time as well as its intensity ratio must be calculated on the basis of the phase shifts and the overall amplitudes at both modulation frequencies. Following the mathematical descriptions in section 3.3, the solution of the particular measurement point is clearly defined by both parameters, as the junction of the intersection lines between two continuously and always monotonous planes is determined explicitly. Consequently, the intersection lines of the reference matrix and the actual values, which is spanned as three-dimensional plane, have to be computed. Subsequently, the junction of both intersection lines will

yield the actual oxygen concentration and the actual temperature. The estimated error propagated across the evaluation process originated form the measurement uncertainty in the recorded phase shifts can be reduced by a more elaborated measurement setup, including an enhanced signal-to-noise ratio and a greater dynamics of the overall phase shifts. For a description of the evaluation approach in more detail, reference is made to section 3.3.
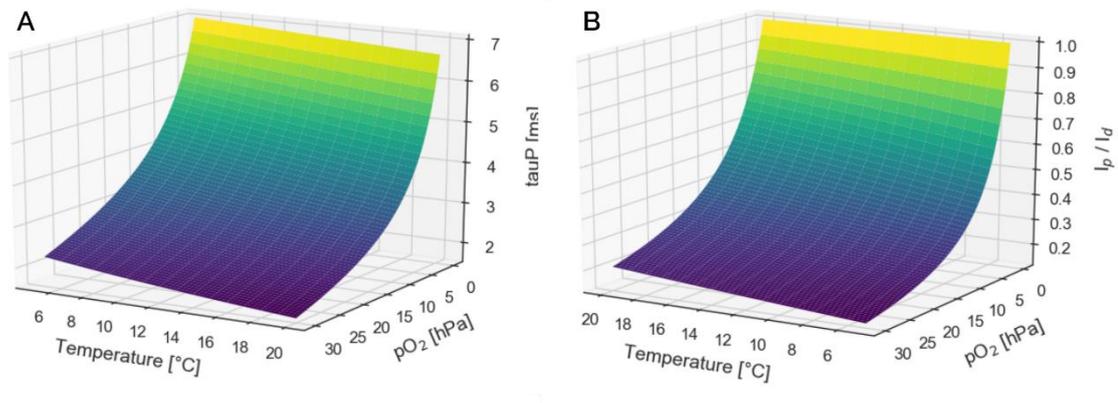
Figure 9-3: Extrapolation of the decay time (**A**) and the intensity ratio (**B**) of a reference matrix based on the conducted calibration measurements. A linear behavior was assumed for the temperature dependency, whereas the oxygen dependency is best described by the two-site-model.

To sum up the preliminary experiments, the designed single-layered dual optical sensor sensitive to oxygen and temperature was successfully applied in the gaseous phase. Based on the calibration measurements, the dynamic range of the sensor was determined within an oxygen range of 0-30 hPa, whereas the sensor can be used at least within a temperature range of 5-20 °C. Further projects will characterize the maximal measurement range of the hybrid sensor and will apply the sensor to the aqueous phase for continuous monitoring.

Part IV

Future prospects and conclusions

10. Summary and conclusion

The purpose of the dissertation was to develop an optical sensor feasible for multiparametric sensing in coastal environments. Initially, the project was divided into two individual parts: one was dedicated to the reliable identification of potentially harmful phytoplankton at unicellular level, whereas the second part was dedicated to the evaluation of multiplexed optical sensors, their appropriate calibration as well as to the propagation of errors during evaluation.

Regarding the first project, a miniaturized multi-channel phytoplankton sensor for applications in coastal environments was successfully developed and tested in natural conditions. The sensor enables both, a semi-quantitative and a qualitative analysis, though the quantitative analysis was proven only once, while the qualitative analysis and reliable identification of potentially toxic algae was optimized and implemented into a software interface. The identification of the phytoplankton sample bases on the inherent pigment composition of the sample using the linear discriminant analysis as multivariate pattern recognition algorithm. This algorithm enables the separation of the phytoplankton with respect to a well characterized reference database and the subsequent classification due to the Mahalanobis distance. In general, the evaluation procedure and the sensor are designed to identify potentially toxic phytoplankton samples, in particular cyanobacteria and dinoflagellates, at low concentration levels. To enhance the sensor performance in terms of inter- and intra-comparability as well as repeatability, internal calibration and standardization strategies were installed using rhodamine-101 as quantum counter.

For real-world applications, the performance of the sensor as well as its possibility for linkage into a broader sensor observation network were successfully demonstrated during multiple international field trips. In the port of Genoa (Italy), the sensor was integrated into a submersible housing combining different sensors for marine monitoring and deployed to demonstrate the long-term stability of the sensor. In Arcachon (France) and Gironde (France), the sensor was again integrated into a submersible housing for comprehensive analysis on a high-resolution scale in space and time. In all cases, the ease of use of the sensor and its versatility was demonstrated as well as its separation and identification efficiency. During the dissertation, the algae sensor thus reached the technical readiness level of 7 which correspond to a system that is adequately validated in space. Currently, the sensor is involved into the research project 'JERICO-next' connecting multiple coastal monitoring tools in one 'free of charge' transnational observation network.

Despite all the advantages mentioned, the ALPACA has some limitations concerning its application. Currently, the sensor is only suitable for coastal environments and for the analysis of surface water as the pigmentation of the phytoplankton relies on multiple external (growing) conditions. However, this limitation does not arise from technical restrictions and can easily be overcome by adapting the training database to further well-described phytoplankton samples. Therefore, the

algae sensor might be a valuable contribution to already existing monitoring tools applied in coastal environments. Its versatility, the low purchase price and the minimal servicing required for system maintenance makes it an attractive option also for small agencies, universities and scientific surveys to study phytoplankton dynamics on-site on a high-resolution scale in space and time.

Thus, a natural progression of the algae project is to expand the mathematical algorithm including further parameters for classification, such as the growth stage or the light intensity used for cultivation of the training samples. Additionally, the sensor could be extended on the detector side enabling the combination of multiple optical filters in order to study the emission behavior of the phytoplankton samples and thus to further reduce cross-interferences, such as colored dissolved organic matters or hydrocarbon materials.

Regarding the second project, different simulation and evaluation strategies were designed for simultaneous determination of multiplexed sensors based on the dual lifetime referencing (DLR) technique. Since the modified DLR technique includes multiple assumptions for facilitated evaluation, such as mono-exponential decay of the indicator dyes and non-reciprocal interaction between the individual sensor layers, the system is prone to errors. However, this challenge can be overcome by inclusion of the measurement uncertainty and the investigation of its propagation across the evaluation process. These considerations and the final evaluation strategies were implemented in a software interface enabling the simulation of anticipated phase shifts and the evaluation of multiple measurement data including an estimation of the maximal error. Moreover, preliminary experiments were conducted for the single-layered hybrid sensor feasible for simultaneous determination of oxygen and temperature in the gaseous phase. In theory, the calibration and analysis of the sensor is convenient and straightforward, however the measurement setup suffers from some challenges due to the long luminescence decay time of the indicator.

In conclusion, investigations concerning the measurement setup of these dual optical sensors and their application in the aqueous phase are crucial for the overall aim of a multiparametric sensor network for coastal monitoring. In particular, the trend goes towards sensor observation networks and optical sensors have proven to be valuable in elucidating and enhancing our understanding of chemical and biological processes in marine life. For instance, the dual optical sensor can be expanded by integration of further indicator dyes with similar spectral properties for multiplexed analysis of a multi hybrid sensor. In this case however, it might be challenging to clearly separate the superimposed sensor signal, though an iterative evaluation might be a promising approach. However, this might be an interesting challenge for further projects. Moreover, the hybrid sensors can be integrated into the algae sensor for simultaneous monitoring of the algae population dynamics and related physiological parameters within one miniaturized and low-cost sensor feasible for versatile applications.

Part V

Appendix

11. Graphical user interfaces

In this chapter, both graphical user interfaces for evaluation and classification of the phytoplankton as well as for evaluation of the optical hybrid sensors are provided. The corresponding exe-programs are stored on the USB card slider and can be installed as previously explained in chapter 6 and chapter 8.

12. References

(1) Lee, R. E. *Phycology.*; Cambridge University Press: Leiden, 2008.

(2) McCormick, P. V.; Cairns, J. Algae as Indicators of Environmental Change. *Journal of Applied Phycology* **1994**, *6* (5–6), 509–526. https://doi.org/10.1007/BF02182405.

(3) Hallegraeff, G. M. A Review of Harmful Algal Blooms and Their Apparent Global Increase*. *Phycologia* **1993**, *32* (2), 79–99. https://doi.org/10.2216/i0031-8884-32-2-79.1.

(4) Pitois, S.; Jackson, M. H.; Wood, B. J. B. Problems Associated with the Presence of Cyanobacteria in Recreational and Drinking Waters. *International Journal of Environmental Health Research* **2000**, *10* (3), 203–218. https://doi.org/10.1080/09603120050127158.

(5) Waeber, M.-L. T.; Bakker, E.; Nardin, C.; Mongin, S.; Prado, E.; Botia, M. C.; Mazaikoff, B.; Luxenburger, F.; Klimant, I.; Mistlberger, G.; et al. FP7-OCEAN-2013 - SCHeMA: Integrated in Situ Chemical MApping Probes. In *OCEANS 2015 - Genova*; IEEE: Genova, 2015; pp 1–5. https://doi.org/10.1109/OCEANS-Genova.2015.7271560.

(6) Sharma, O. P. *Textbook of Algae*, 20. reprint.; New Delhi [u.a.]: Tata McGraw-Hill, 2008.

(7) Cavalier-Smith, T. A Revised Six-Kingdom System of Life. *Biological Reviews* **2007**, *73* (3), 203–266. https://doi.org/10.1111/j.1469-185X.1998.tb00030.x.

(8) Reichenbach, H. G. L. *Handbuch Des Natürlichen Pflanzensystems Nach Allen Seinen Classen, Ordnungen Und Familien: Nebst Naturgemä\s Ser Gruppirung Der Gattungen*; Arnoldi, 1837.

(9) Wettstein, R. *Handbuch der systematischen Botanik*, 3rd ed.; Handbuch der systematischen Botanik; F. Deuticke: Wien und Leipzig, 1901.

(10) Anderson, D. M.; Andersen, P.; Bricelj, M. V.; Cullen, J. J.; Rensel, J. J. E. *Monitoring and Management Strategies for Harmful Algal Blooms in Coastal Waters*; APEC Report # 201-MR-01.1 59; Asia Pacific Economic Program and Intergovernmental Oceanographic Commission of UNESCO: France, 2001; p 268.

(11) Masó, M.; Garcés, E. Harmful Microalgae Blooms (HAB); Problematic and Conditions That Induce Them. *Marine Pollution Bulletin* **2006**, *53* (10–12), 620–630. https://doi.org/10.1016/j.marpolbul.2006.08.006.

(12) Fichez, R.; Jickells, T. D.; Edmunds, H. M. Algal Blooms in High Turbidity, a Result of the Conflicting Consequences of Turbulence on Nutrient Cycling in a Shallow Water Estuary. *Estuarine, Coastal and Shelf Science* **1992**, *35* (6), 577–592. https://doi.org/10.1016/S0272-7714(05)80040-X.

(13) Hoagland, P.; Anderson, D. M.; Kaoru, Y.; White, A. W. The Economic Effects of Harmful Algal Blooms in the United States: Estimates, Assessment Issues, and Information Needs. *Estuaries* **2002**, *25* (4), 819–837. https://doi.org/10.1007/BF02804908.

(14) Malapascua, J.; Jerez, C.; Sergejevová, M.; Figueroa, F.; Masojídek, J. Photosynthesis Monitoring to Optimize Growth of Microalgal Mass Cultures: Application of Chlorophyll Fluorescence Techniques. *Aquatic Biology* **2014**, *22*, 123–140. https://doi.org/10.3354/ab00597.

(15) Zielinski, O.; Busch, J. A.; Cembella, A. D.; Daly, K. L.; Engelbrektsson, J.; Hannides, A. K.; Schmidt, H. Detecting Marine Hazardous Substances and Organisms: Sensors for Pollutants, Toxins, and Pathogens. *Ocean Science Discussions* **2009**, *6* (2), 953–1005. https://doi.org/10.5194/osd-6-953-2009.

(16) Anderson, D. M. Approaches to Monitoring, Control and Management of Harmful Algal Blooms (HABs). *Ocean & Coastal Management* **2009**, *52* (7), 342–347. https://doi.org/10.1016/j.ocecoaman.2009.04.006.

(17) Sellner, K. G.; Doucette, G. J.; Kirkpatrick, G. J. Harmful Algal Blooms: Causes, Impacts and Detection. *Journal of Industrial Microbiology and Biotechnology* **2003**, *30* (7), 383–406. https://doi.org/10.1007/s10295-003-0074-9.

(18)   Sparnocchia, S.; Nair, R.; Petihakis, G.; Aydoğdu, A.; Dobricic, S.; Farcy, P.; Martinelli, M.; Petersen, W.; Petit de la Villeon, L. An Interlinked Coastal Observatory Network for Europe. *Journal of Operational Oceanography* **2016**, *9* (sup1), s193–s201. https://doi.org/10.1080/1755876X.2015.1114808.

(19)   Moore, C.; Barnard, A.; Fietzek, P.; Lewis, M. R.; Sosik, H. M.; White, S.; Zielinski, O. Optical Tools for Ocean Monitoring and Research. *Ocean Science* **2009**, *5* (4), 661–684. https://doi.org/10.5194/os-5-661-2009.

(20)   Schofield, O.; Grzymski, J.; Bissett, W. P.; Kirkpatrick, G. J.; Millie, D. F.; Moline, M.; Roesler, C. S. OPTICAL MONITORING AND FORECASTING SYSTEMS FOR HARMFUL ALGAL BLOOMS: POSSIBILITY OR PIPE DREAM? *Journal of Phycology* **1999**, *35* (6), 1477–1496. https://doi.org/10.1046/j.1529-8817.1999.3561477.x.

(21)   Dickey, T.; Lewis, M.; Chang, G. Optical Oceanography: Recent Advances and Future Directions Using Global Remote Sensing and in Situ Observations. *Reviews of Geophysics* **2006**, *44* (1). https://doi.org/10.1029/2003RG000148.

(22)   Manoylov, K. M. Taxonomic Identification of Algae (Morphological and Molecular): Species Concepts, Methodologies, and Their Implications for Ecological Bioassessment. *Journal of Phycology* **2014**, *50* (3), 409–424. https://doi.org/10.1111/jpy.12183.

(23)   Bidigare, R. R.; Ondrusek, M. E.; Morrow, J. H.; Kiefer, D. A. In-Vivo Absorption Properties of Algal Pigments. In *Proceedings of SPIE. Ocean Optics X*; 1990; Vol. 1302, p 290. https://doi.org/10.1117/12.21451.

(24)   Cullen, J. J.; Ciotti, Á. M.; Davis, R. F.; Lewis, M. R. Optical Detection and Assessment of Algal Blooms. *Limnology and Oceanography* **1997**, *42* (5part2), 1223–1239. https://doi.org/10.4319/lo.1997.42.5_part_2.1223.

(25)   Millie, D. F.; Schofield, O. M. ??E.; Kirkpatrick, G. J.; Johnsen, G.; Evens, T. J. Using Absorbance and Fluorescence Spectra to Discriminate Microalgae. *European Journal of Phycology* **2002**, *37* (3), 313–322. https://doi.org/10.1017/S0967026202003700.

(26)   Bricaud, A.; Stramski, D. Spectral Absorption Coefficients of Living Phytoplankton and Nonalgal Biogenous Matter: A Comparison between the Peru Upwelling Areaand the Sargasso Sea. *Limnology and Oceanography* **1990**, *35* (3), 562–582. https://doi.org/10.4319/lo.1990.35.3.0562.

(27)   Giordano, M.; Ratti, S.; Domenighini, A.; Vogt, F. Spectroscopic Classification of 14 Different Microalga Species: First Steps towards Spectroscopic Measurement of Phytoplankton Biodiversity. *Plant Ecology & Diversity* **2009**, *2* (2), 155–164. https://doi.org/10.1080/17550870903353088.

(28)   Haraguchi, L.; Jakobsen, H. H.; Lundholm, N.; Carstensen, J. Phytoplankton Community Dynamic: A Driver for Ciliate Trophic Strategies. *Frontiers in Marine Science* **2018**, *5*. https://doi.org/10.3389/fmars.2018.00272.

(29)   Olson, R. J.; Sosik, H. M. A Submersible Imaging-in-Flow Instrument to Analyze Nano-and Microplankton: Imaging FlowCytobot: In Situ Imaging of Nano- and Microplankton. *Limnology and Oceanography: Methods* **2007**, *5* (6), 195–203. https://doi.org/10.4319/lom.2007.5.195.

(30)   Sosik, H. M.; Olson, R. J.; Neubert, M. G.; Shalapyonok, A.; Solow, A. R. Growth Rates of Coastal Phytoplankton from Time-Series Measurements with a Submersible Flow Cytometer. *Limnology and Oceanography* **2003**, *48* (5), 1756–1765. https://doi.org/10.4319/lo.2003.48.5.1756.

(31)   Dubelaar, G. B. J.; Jonker, R. R. Flow Cytometry as a Tool for the Study of Phytoplankton*. *Scientia Marina* **2000**, *64* (2), 135–156.

(32)   Erickson, J. S.; Hashemi, N.; Sullivan, J. M.; Weidemann, A. D.; Ligler, F. S. *In Situ* Phytoplankton Analysis: There's Plenty of Room at the Bottom. *Analytical Chemistry* **2012**, *84* (2), 839–850. https://doi.org/10.1021/ac201623k.

(33)   Ateya, D. A.; Erickson, J. S.; Howell, P. B.; Hilliard, L. R.; Golden, J. P.; Ligler, F. S. The Good, the Bad, and the Tiny: A Review of Microflow Cytometry. *Analytical and Bioanalytical Chemistry* **2008**, *391* (5), 1485–1498. https://doi.org/10.1007/s00216-007-1827-5.

(34)   Golden, J. P.; Hashemi, N.; Erickson, J. S.; Ligler, F. S. A Microflow Cytometer for Optical Analysis of Phytoplankton; Miller, B. L., Fauchet, P. M., Eds.; 2012; p 82120G. https://doi.org/10.1117/12.905679.

(35)   *Applications of in Situ Fluorometers in Nearshore Water*; Alliance for Coastal Technologies, Series Ed.; Monograph or Serial Issue; Workshop Proceedings ACT 05-03, UMCES CBL 05-078; Alliance for Coastal Technologies, NOAA: Solomons, MD, 2005.

(36)   Millie, D. F.; Schofield, O. M.; Kirkpatrick, G. J.; Johnsen, G.; Tester, P. A.; Vinyard, B. T. Detection of Harmful Algal Blooms Using Photopigments and Absorption Signatures: A Case Study of the Florida Red Tide Dinoflagellate, *Gymnodinium Breve*. *Limnology and Oceanography* **1997**, *42* (5part2), 1240–1251. https://doi.org/10.4319/lo.1997.42.5_part_2.1240.

(37)   Minu, P.; Shaju, S. S.; Muhamed Ashraf, P.; Meenakumari, B. Phytoplankton Community Characteristics in the Coastal Waters of the Southeastern Arabian Sea. *Acta Oceanologica Sinica* **2014**, *33* (12), 170–179. https://doi.org/10.1007/s13131-014-0571-x.

(38)   Poryvkina, L.; Babichenko, S.; Kaitala, S.; Kuosa, H.; Shalapjonok, A. Spectral Fluorescence Signatures in the Characterization of Phytoplankton Community Composition. *Journal of Plankton Research* **1994**, *16* (10), 1315–1327. https://doi.org/10.1093/plankt/16.10.1315.

(39)   Schopfer, P.; Brennicke, A.; Mohr, H. *Pflanzenphysiologie*, 7. Aufl.; Spektrum, Akad. Verl: Heidelberg, 2010.

(40)   Liberton, M.; Page, L. E.; O'Dell, W. B.; O'Neill, H.; Mamontov, E.; Urban, V. S.; Pakrasi, H. B. Organization and Flexibility of Cyanobacterial Thylakoid Membranes Examined by Neutron Scattering. *Journal of Biological Chemistry* **2013**, *288* (5), 3632–3640. https://doi.org/10.1074/jbc.M112.416933.

(41)   Govindjee; Amesz, J.; Fork, D. C. *Light Emission by Plants and Bacteria*; Cell biology - A series of monographs; Academic Press, Inc., 1986.

(42)   Johnson, M. P. Photosynthesis. *Essays In Biochemistry* **2016**, *60* (3), 255–273. https://doi.org/10.1042/EBC20160016.

(43)   Croce, R.; van Amerongen, H. Natural Strategies for Photosynthetic Light Harvesting. *Nature Chemical Biology* **2014**, *10* (7), 492–501. https://doi.org/10.1038/nchembio.1555.

(44)   Valeur, B. *Molecular Fluorescence Principles and Applications*; Wiley-VCH: Weinheim, 2002.

(45)   Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer: New York, 2006.

(46)   Duysens, L. N. M. Introduction to (Bacterio)Chlorophyll Emission: A Historical Perspective. In *Light emission by plants and bacteria*; Cell biology - A series of monographs; Academic Press, Inc., 1986.

(47)   Gust, D.; Moore, T. A.; Moore, A. L. Molecular Mimicry of Photosynthetic Energy and Electron Transfer. *Accounts of Chemical Research* **1993**, *26* (4), 198–205. https://doi.org/10.1021/ar00028a010.

(48)   van Grondelle, R.; Dekker, J. P.; Gillbro, T.; Sundstrom, V. Energy Transfer and Trapping in Photosynthesis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1994**, *1187* (1), 1–65. https://doi.org/10.1016/0005-2728(94)90166-X.

(49)   Dexter, D. L. A Theory of Sensitized Luminescence in Solids. *The Journal of Chemical Physics* **1953**, *21* (5), 836–850. https://doi.org/10.1063/1.1699044.

(50)   Wilhelm, C.; Manns, L. Changes in Pigmentation of Phytoplankton Species during Growth and Stationary Phase — Consequences for Reliability of Pigment-Based Methods of Biomass Determination. *Journal of Applied Phycology* **1991**, *3* (4), 305–310. https://doi.org/10.1007/BF02392883.

(51)   Laviale, M.; Neveux, J. Relationships between Pigment Ratios and Growth Irradiance in 11 Marine Phytoplankton Species. *Marine Ecology Progress Series* **2011**, *425*, 63–77. https://doi.org/10.3354/meps09013.

(52)   Cullen, J. J. The Deep Chlorophyll Maximum: Comparing Vertical Profiles of Chlorophyll *a*. *Canadian Journal of Fisheries and Aquatic Sciences* **1982**, *39* (5), 791–803. https://doi.org/10.1139/f82-108.

(53)   Lazár, D. Chlorophyll a Fluorescence Induction1Dedicated to Docent Jan Nauš on the Occasion of His 50th Birthday.1. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1999**, *1412* (1), 1–28. https://doi.org/10.1016/S0005-2728(99)00047-X.

(54)   Stirbet, A.; Govindjee. On the Relation between the Kautsky Effect (Chlorophyll a Fluorescence Induction) and Photosystem II: Basics and Applications of the OJIP Fluorescence Transient. *Journal of Photochemistry and Photobiology B: Biology* **2011**, *104* (1–2), 236–257. https://doi.org/10.1016/j.jphotobiol.2010.12.010.

(55)   Kautsky, H.; Hirsch, A. Neue Versuche zur Kohlensäureassimilation. *Die Naturwissenschaften* **1931**, *19* (48), 964–964. https://doi.org/10.1007/BF01516164.

(56)   Paerl, H. W.; Valdes, L. M.; Pinckney, J. L.; Piehler, M. F.; Dyble, J.; Moisander, P. H. Phytoplankton Photopigments as Indicators of Estuarine and Coastal Eutrophication. *BioScience* **2003**, *53* (10), 953. https://doi.org/10.1641/0006-3568(2003)053[0953:PPAIOE]2.0.CO;2.

(57)   Entry, J. A.; Gottlieb, A.; Jayachandran, K. V.; Ogram, A. *Microbiology of the Everglades Ecosystem*; 2015.

(58)   MacIntyre, H. L.; Lawrenz, E.; Richardson, T. L. Taxonomic Discrimination of Phytoplankton by Spectral Fluorescence. In *Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications*; Suggett, D. J., Prášil, O., Borowitzka, M. A., Eds.; Springer Netherlands: Dordrecht, 2010; pp 129–169.

(59)   Mackey, M.; Mackey, D.; Higgins, H.; Wright, S. CHEMTAX - a Program for Estimating Class Abundances from Chemical Markers:Application to HPLC Measurements of Phytoplankton. *Marine Ecology Progress Series* **1996**, *144*, 265–283. https://doi.org/10.3354/meps144265.

(60)   *Algal Cultures, Analogues of Blooms and Applications*; Subba Rao, D. V., Ed.; Science Publishers: Enfield, NH, 2006.

(61)   Gieskes, W. W. C.; Kraay, G. W.; Nontji, A.; Setiapermana, D.; Sutomo. Monsoonal Alternation of a Mixed and a Layered Structure in the Phytoplankton of the Euphotic Zone of the Banda Sea (Indonesia): A Mathematical Analysis of Algal Pigment Fingerprints. *Netherlands Journal of Sea Research* **1988**, *22* (2), 123–137. https://doi.org/10.1016/0077-7579(88)90016-6.

(62)   Everitt, D. A.; Wright, S. W.; Volkman, J. K.; Thomas, D. P.; Lindstrom, E. J. Phytoplankton Community Compositions in the Western Equatorial Pacific Determined from Chlorophyll and Carotenoid Pigment Distributions. *Deep Sea Research Part A. Oceanographic Research Papers* **1990**, *37* (6), 975–997. https://doi.org/10.1016/0198-0149(90)90106-6.

(63)   Letelier, R. M.; Bidigare, R. R.; Hebel, D. V.; Ondrusek, M.; Winn, C. D.; Karl, D. M. Temporal Variability of Phytoplankton Community Structure Based on Pigment Analysis. *Limnology and Oceanography* **1993**, *38* (7), 1420–1437. https://doi.org/10.4319/lo.1993.38.7.1420.

(64)   *Phytoplankton Pigments: Characterization, Chemotaxonomy and Applications in Oceanography*; Roy, S., Llewellyn, C., Egeland, E. S., Johnsen, G., Eds.; Cambridge University Press: Cambridge, 2011. https://doi.org/10.1017/CBO9780511732263.

(65)   Beutler, M.; Wiltshire, K. H.; Meyer, B.; Moldaenke, C.; Lüring, C.; Meyerhöfer, M.; Hansen, U.-P.; Dau, H. A Fluorometric Method for the Differentiation of Algal Populations in Vivo and in Situ. *Photosynthesis Research* **2002**, *72* (1), 39–53. https://doi.org/10.1023/A:1016026607048.

(66)   Yentsch, C. S.; Phinney, D. A. Fluorescence Spectral Signatures for Studies of Marine Phytoplankton. In *Mapping Strategies in Chemical Oceanography*; Zirino, A., Ed.; American Chemical Society: Washington, D.C., 1985; Vol. 209, pp 259–274. https://doi.org/10.1021/ba-1985-0209.ch013.

(67)   Hodges, M.; Barber, J. State 1-State 2 Transitions in a Unicellular Green Algae : Analysis of In Vivo Chlorophyll Fluorescence Induction Curves in the Presence of 3-(3,4-Dichlorophenyl)-1, 1-Dimethylurea (DCMU). *PLANT PHYSIOLOGY* **1983**, *72* (4), 1119–1122. https://doi.org/10.1104/pp.72.4.1119.

(68)   Vincent, W. F. Fluorescence Properties of the Freshwater Phytoplankton: Three Algal Classes Compared. *British Phycological Journal* **1983**, *18* (1), 5–21. https://doi.org/10.1080/00071618300650021.

(69)   Bryant, D. A. Phycoerythrocyanin and Phycoerythrin: Properties and Occurrence in Cyanobacteria. *Microbiology* **1982**, *128* (4), 835–844. https://doi.org/10.1099/00221287-128-4-835.

(70)   Higgins, H. W.; Wright, S. W.; Schlüter, L. Quantitative Interpretation of Chemotaxonomic Pigment Data. In *Phytoplankton Pigments*; Roy, S., Llewellyn, C., Egeland, E. S., Johnsen, G., Eds.; Cambridge University Press: Cambridge, 2011; pp 257–313. https://doi.org/10.1017/CBO9780511732263.010.

(71)   *Marine Biogenic Lipids, Fats, and Oils*; Ackman, R. G., Ed.; CRC Press: Boca Raton, Fla, 1989.

(72)   Bricaud, A. Natural Variability of Phytoplanktonic Absorption in Oceanic Waters: Influence of the Size Structure of Algal Populations. *Journal of Geophysical Research* **2004**, *109* (C11). https://doi.org/10.1029/2004JC002419.

(73)   Bjørnland, T.; Aguilar-Martinez, M. Carotenoids in Red Algae. *Phytochemistry* **1976**, *15* (2), 291–296. https://doi.org/10.1016/S0031-9422(00)89006-8.

(74)   Whittle, S. J.; Casselton, P. J. The Chloroplast Pigments of the Algal Classes Eustigmatophyceae and Xanthophyceae. II. Xanthophyceae. *British Phycological Journal* **1975**, *10* (2), 192–204. https://doi.org/10.1080/00071617500650181.

(75)   Zapata, M.; Jeffrey, S.; Wright, S.; Rodríguez, F.; Garrido, J.; Clementson, L. Photosynthetic Pigments in 37 Species (65 Strains) of Haptophyta: Implications for Oceanography and Chemotaxonomy. *Marine Ecology Progress Series* **2004**, *270*, 83–102. https://doi.org/10.3354/meps270083.

(76)   Zapata, M.; Fraga, S.; Rodríguez, F.; Garrido, J. Pigment-Based Chloroplast Types in Dinoflagellates. *Marine Ecology Progress Series* **2012**, *465*, 33–52. https://doi.org/10.3354/meps09879.

(77)   Zapatal, M.; Rodriguezl, F.; Garrido, J. L. Separation of Chlorophylls and Carotenoids from Marine Phytoplankton: A New HPLC Method Using a Reversed Phase C8 Column and Pyridine-Containing Mobile Phases. *Mar Ecol Prog Ser* **2000**, No. 195, 29–45.

(78)   Owens, T. G.; Gallagher, J. C.; Alberte, R. S. PHOTOSYNTHETIC LIGHT-HARVESTING FUNCTION OF VIOLAXANTHIN IN NANNOCHLOROPSIS SPP. (EUSTIGMATOPHYCEAE)1. *Journal of Phycology* **2004**, *23* (1), 79–85. https://doi.org/10.1111/j.0022-3646.1987.00079.x.

(79)   *Real-Time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling*; Babin, M., Roesler, C. S., Cullen, J. J., UNESCO, Eds.; Oceanographic methodology series; UNESCO Publishing: Paris, 2008.

(80)   Miller, J. N.; Miller, J. C. Discriminant Analysis. In *Statistics and Chemometrics for Analytical Chemistry*; Pearson Education Limited: Harlow, England, 2005; p 268.

(81)   Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*, 5. ed., [Nachdr.].; Pearson Prentice Hall: Harlow, England, 2009.

(82)   *Chemometrik: Grundlagen und Anwendungen*; Danzer, K., Hobert, H., Fischbacher, C., Jagemann, K.-U., Eds.; Springer: Berlin, 2001.

(83)   Pechenizkiy, M.; Tsymbal, A.; Puuronen, S. On Combining Principal Components with Fisher's Linear Discriminants for Supervised Learning. *Foundations of Computing and Decision Sciences* **2006**, *31* (1), 59–73.

(84)   De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D. L. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems* **2000**, *50* (1), 1–18. https://doi.org/10.1016/S0169-7439(99)00047-7.

(85)   Minnhagen, S.; Carvalho, W. F.; Salomon, P. S.; Janson, S. Chloroplast DNA Content in *Dinophysis* (*Dinophyceae*) from Different Cell Cycle Stages Is Consistent with Kleptoplasty. *Environmental Microbiology* **2008**, *10* (9), 2411–2417. https://doi.org/10.1111/j.1462-2920.2008.01666.x.

(86)   Gansert, D.; Arnold, M.; Borisov, S.; Krause, C.; Müller, A.; Stangelmayer, A.; Wolfbeis, O. Hybrid Optodes (HYBOP). In *Biophotonics*; Popp, J., Strehle, M., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, FRG, 2006; pp 477–518. https://doi.org/10.1002/3527608842.ch9.

(87)   McDonagh, C.; Burke, C. S.; MacCraith, B. D. Optical Chemical Sensors. *Chemical Reviews* **2008**, *108* (2), 400–422. https://doi.org/10.1021/cr068102g.

(88)   Borisov, S. M.; Seifner, R.; Klimant, I. A Novel Planar Optical Sensor for Simultaneous Monitoring of Oxygen, Carbon Dioxide, PH and Temperature. *Analytical and Bioanalytical Chemistry* **2011**, *400* (8), 2463–2474. https://doi.org/10.1007/s00216-010-4617-4.

(89)   Stich, M. I. J.; Fischer, L. H.; Wolfbeis, O. S. Multiple Fluorescent Chemical Sensing and Imaging. *Chemical Society Reviews* **2010**, *39* (8), 3102. https://doi.org/10.1039/b909635n.

(90)   Nagl, S.; Wolfbeis, O. S. Optical Multiple Chemical Sensing: Status and Current Challenges. *The Analyst* **2007**, *132* (6), 507. https://doi.org/10.1039/b702753b.

(91)   *Sensor Systems for Environmental Monitoring Volume One: Sensor Technologies*; Springer Netherlands: Dordrecht, 1996.

(92)   Course of the International School of Quantum Electronics on Optical Sensors and Microsystems: New Concepts, M., Technologies; Martellucci, S.; Chester, A. N.; Mignani, A. G. *Optical Sensors and Microsystems: New Concepts, Materials, Technologies*; Kluwer Academic/Plenum Publishers: New York, 2000.

(93)   Wolfbeis, O. S. Materials for Fluorescence-Based Optical Chemical Sensors. *Journal of Materials Chemistry* **2005**, *15* (27–28), 2657. https://doi.org/10.1039/b501536g.

(94)   Borisov, S. M.; Vasylevska, A. S.; Krause, Ch.; Wolfbeis, O. S. Composite Luminescent Material for Dual Sensing of Oxygen and Temperature. *Advanced Functional Materials* **2006**, *16* (12), 1536–1542. https://doi.org/10.1002/adfm.200500778.

(95)   Borisov, S. M.; Klimant, I. New Luminescent Oxygen-Sensing and Temperature-Sensing Materials Based on Gadolinium(III) and Europium(III) Complexes Embedded in an Acridone–Polystyrene Conjugate. *Analytical and Bioanalytical Chemistry* **2012**, *404* (10), 2797–2806. https://doi.org/10.1007/s00216-012-6244-8.

(96)   Schröder, C. R.; Polerecky, L.; Klimant, I. Time-Resolved PH/PO $_2$ Mapping with Luminescent Hybrid Sensors. *Analytical Chemistry* **2007**, *79* (1), 60–70. https://doi.org/10.1021/ac0606047.

(97)   Schroeder, C. R.; Neurauter, G.; Klimant, I. Luminescent Dual Sensor for Time-Resolved Imaging of PCO2 and PO2 in Aquatic Systems. *Microchimica Acta* **2007**, *158* (3–4), 205–218. https://doi.org/10.1007/s00604-006-0696-5.

(98)   Harvey, D. *Modern Analytical Chemistry*; McGraw-Hill: Boston, 2000.

(99)   Klimant, I.; Huber, C.; Liebsch, G.; Wolfbeis, O. S. Dual Lifetime Referencing (ĺDLRĺ) - A New Scheme for Converting Fluorescence Intensity into a Frequency-Domain or Time-Domain Information. In *New Trends in Fluorescence Spectroscopy: Applications to Chemical and Life Sciences*; Springer, 2001; pp 257–274.

(100)   Borisov, S. M.; Neurauter, G.; Schroeder, C.; Klimant, I.; Wolfbeis, O. S. Modified Dual Lifetime Referencing Method for Simultaneous Optical Determination and Sensing of Two Analytes. *Applied Spectroscopy* **2006**, *60* (10), 1167–1173. https://doi.org/10.1366/000370206778664590.

(101)   Huber, C.; Klimant, I.; Krause, C.; Wolfbeis, O. S. Dual Lifetime Referencing as Applied to a Chloride Optical Sensor. *Analytical Chemistry* **2001**, *73* (9), 2097–2103. https://doi.org/10.1021/ac9914364.

(102)   Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, Third edition.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2017.

(103)   *Taschenbuch der Mathematik*, 6., vollst. überarb. und erg. Aufl., Nachdruck.; Bronštejn, I. N., Semendjaev, K. A., Eds.; Deutsch: Frankfurt am Main, 2006.

(104)   Steinegger, A.; Klimant, I.; Borisov, S. M. Purely Organic Dyes with Thermally Activated Delayed Fluorescence-A Versatile Class of Indicators for Optical Temperature Sensing. *Advanced Optical Materials* **2017**, *5* (18), 1700372. https://doi.org/10.1002/adom.201700372.

(105)   Demas, J. N.; DeGraff, B. A.; Xu, Wenying. Modeling of Luminescence Quenching-Based Sensors: Comparison of Multisite and Nonlinear Gas Solubility Models. *Analytical Chemistry* **1995**, *67* (8), 1377–1380. https://doi.org/10.1021/ac00104a012.

(106)   Liebsch, G.; Klimant, I.; Wolfbeis, O. S. Luminescence Lifetime Temperature Sensing Based on Sol-Gels and Poly(Acrylonitrile)s Dyed with Ruthenium Metal-Ligand Complexes. *Advanced*

Materials **1999**, *11* (15), 1296–1299. https://doi.org/10.1002/(SICI)1521-4095(199910)11:15<1296::AID-ADMA1296>3.0.CO;2-B.

(107)   Staudinger, C.; Breininger, J.; Klimant, I.; Borisov, S. M. Near-Infrared Fluorescent Aza-BODIPY Dyes for Sensing and Imaging of PH from the Neutral to Highly Alkaline Range. *The Analyst* **2019**, *144* (7), 2393–2402. https://doi.org/10.1039/C9AN00118B.

(108)   Staudinger, C.; Strobl, M.; Fischer, J. P.; Thar, R.; Mayr, T.; Aigner, D.; Müller, B. J.; Müller, B.; Lehner, P.; Mistlberger, G.; et al. A Versatile Optode System for Oxygen, Carbon Dioxide, and PH Measurements in Seawater with Integrated Battery and Logger: *A Versatile Optode System for O $_2$, CO $_2$, and PH* . *Limnology and Oceanography: Methods* **2018**, *16* (7), 459–473. https://doi.org/10.1002/lom3.10260.

(109)   Wang, X.; Wolfbeis, O. S.; Meier, R. J. Luminescent Probes and Sensors for Temperature. *Chemical Society Reviews* **2013**, *42* (19), 7834. https://doi.org/10.1039/c3cs60102a.

(110)   Borisov, S. M.; Wolfbeis, O. S. Temperature-Sensitive Europium(III) Probes and Their Use for Simultaneous Luminescent Sensing of Temperature and Oxygen. *Analytical Chemistry* **2006**, *78* (14), 5094–5101. https://doi.org/10.1021/ac060311d.

(111)   Akter, M.; Auguste, M.; Bir, J.; Blanco, E.; Briaudeau, T.; Dapueto, G.; de Cerio, D.; Coccoli, C.; Creemers, M.; Espino, M.; et al. SCHeMA EU Project Summer School Report (Bilbao June 16 – 17, 2016). *CHIMIA International Journal for Chemistry* **2017**, *71* (9), 607–610. https://doi.org/10.2533/chimia.2017.607.

(112)   Glibert, P.; Anderson, D.; Gentien, P.; Granéli, E.; Sellner, K. The Global, Complex Phenomena of Harmful Algal Blooms. *Oceanography* **2005**, *18* (2), 136–147. https://doi.org/10.5670/oceanog.2005.49.

(113)   Bláha, L.; Babica, P.; Maršálek, B. Toxins Produced in Cyanobacterial Water Blooms - Toxicity and Risks. *Interdisciplinary Toxicology* **2009**, *2* (2). https://doi.org/10.2478/v10102-009-0006-2.

(114)   Voloshko, L.; Kopecky, J.; Safronova, T.; Pljusch, A.; Titova, N.; Hrouzek, P.; Drabkova, V. Toxins and Other Bioactive Compounds Produced by Cyanobacteria in Lake Ladoga. *Estonian Journal of Ecology* **2008**, *57* (2), 100. https://doi.org/10.3176/eco.2008.2.02.

(115)   *Guidelines for Safe Recreational Water Environments*; World Health Organization, Ed.; World Health Organization: Geneva, 2003; Vol. 1.

(116)   Dolah, F. M. V. Marine Algal Toxins: Origins, Health Effects, and Their Increased Occurrence. *Environmental Health Perspectives* **2000**, *108(Suppl 1)*, 133–141.

(117)   Nehring, S. Establishment of Thermophilic Phytoplankton Species in the North Sea: Biological Indicators of Climatic Changes? Short Communication. *ICES Journal of Marine Science* **1998**, *55* (4), 818–823. https://doi.org/10.1006/jmsc.1998.0389.

(118)   Beaugrand, G. Reorganization of North Atlantic Marine Copepod Biodiversity and Climate. *Science* **2002**, *296* (5573), 1692–1694. https://doi.org/10.1126/science.1071329.

(119)   Catherine, A.; Escoffier, N.; Belhocine, A.; Nasri, A. B.; Hamlaoui, S.; Yéprémian, C.; Bernard, C.; Troussellier, M. On the Use of the FluoroProbe®, a Phytoplankton Quantification Method Based on Fluorescence Excitation Spectra for Large-Scale Surveys of Lakes and Reservoirs. *Water Research* **2012**, *46* (6), 1771–1784. https://doi.org/10.1016/j.watres.2011.12.056.

(120)   Cooper, E. D.; Bentlage, B.; Gibbons, T. R.; Bachvaroff, T. R.; Delwiche, C. F. Metatranscriptome Profiling of a Harmful Algal Bloom. *Harmful Algae* **2014**, *37*, 75–83. https://doi.org/10.1016/j.hal.2014.04.016.

(121)   Odermatt, D.; Pomati, F.; Pitarch, J.; Carpenter, J.; Kawka, M.; Schaepman, M.; Wüest, A. MERIS Observations of Phytoplankton Blooms in a Stratified Eutrophic Lake. *Remote Sensing of Environment* **2012**, *126*, 232–239. https://doi.org/10.1016/j.rse.2012.08.031.

(122)   Schaap, A.; Rohrlack, T.; Bellouard, Y. Lab on a Chip Technologies for Algae Detection: A Review. *Journal of Biophotonics* **2012**, *5* (8–9), 661–672. https://doi.org/10.1002/jbio.201200051.

(123)   Leeuw, T.; Boss, E.; Wright, D. In Situ Measurements of Phytoplankton Fluorescence Using Low Cost Electronics. *Sensors* **2013**, *13* (6), 7872–7883. https://doi.org/10.3390/s130607872.

(124)   Zeng, L.; Li, D. Development of *In Situ* Sensors for Chlorophyll Concentration Measurement. *Journal of Sensors* **2015**, *2015*, 1–16. https://doi.org/10.1155/2015/903509.

(125)   Yoshida, M.; Horiuchi, T.; Nagasawa, Y. In Situ Multi-Excitation Chlorophyll Fluorometer for Phytoplankton Measurements: Technologies and Applications beyond Conventional Fluorometers; IEEE, 2011; pp 1–4. https://doi.org/10.23919/OCEANS.2011.6107049.

(126)   Hashemi, N.; Erickson, J. S.; Golden, J. P.; Ligler, F. S. Optofluidic Characterization of Marine Algae Using a Microflow Cytometer. *Biomicrofluidics* **2011**, *5* (3), 032009-032009–9. https://doi.org/10.1063/1.3608136.

(127)   Carr, M. R.; Tarran, G. A.; Burkill, P. H. Discrimination of Marine Phytoplankton Species through the Statistical Analysis of Their Flow Cytometric Signatures. *Journal of Plankton Research* **1996**, *18* (7), 1225–1238. https://doi.org/10.1093/plankt/18.7.1225.

(128)   Phinney, D. A.; Cucci, T. L. Flow Cytometry and Phytoplankton. *Cytometry* **1989**, *10* (5), 511–521. https://doi.org/10.1002/cyto.990100506.

(129)   Ramette, A. Multivariate Analyses in Microbial Ecology. *FEMS Microbiology Ecology* **2007**, *62* (2), 142–160. https://doi.org/10.1111/j.1574-6941.2007.00375.x.

(130)   Duan, Y.; Su, R.; Shi, X.; Wang, X.; Zhu, C.; Sun, Y. Fluorescence Discrimination and Determination Technique for Phytoplankton Composition by Coif2 Wavelet Packet. *Journal of Ocean University of China* **2013**, *12* (1), 53–62. https://doi.org/10.1007/s11802-011-1902-9.

(131)   Guillard, R. R. L.; Ryther, J. H. Studies of Marine Planktonic Diatoms; I. Cyclotella Nana Hustedt, and Detonula Confervacea (CLEVE) GRAN. *Canadian Journal of Microbiology* **1962**, *8* (2), 229–239. https://doi.org/10.1139/m62-029.

(132)   Guillard, R. R. L. Culture of Phytoplankton for Feeding Marine Invertebrates. In *Culture of Marine Invertebrate Animals*; Smith, W. L., Chanley, M. H., Eds.; Springer US: Boston, MA, 1975; pp 29–60. https://doi.org/10.1007/978-1-4615-8714-9_3.

(133)   Pilgrim, M. *Dive into Python 3: All You Need to Know to Get off the Ground with Python 3 ; Covers Python 3.1*; The expert's voice in open source; Apress: Berkeley, Calif., 2009.

(134)   Yahaya, S. S. S.; Lim, Y.-F.; Ali, H.; Omar, Z. Robust Linear Discriminant Analysis. *Journal of Mathematics and Statistics* **2016**, *12* (4), 312–316. https://doi.org/10.3844/jmssp.2016.312.316.

(135)   Maxwell, K.; Johnson, G. N. Chlorophyll Fluorescence—a Practical Guide. *Journal of Experimental Botany* **2000**, *51* (345), 659–668. https://doi.org/10.1093/jexbot/51.345.659.

(136)   *Algal Culturing Techniques*; Andersen, R. A., Ed.; Elsevier/Academic Press: Burlington, Mass, 2005.

(137)   Felip, M.; Jordi, C. The Relationship between Phytoplankton Biovolume and Chlorophyll in a Deep Oligotrophic Lake: Decoupling in Their Spatial and Temporal Maxima. *Journal of Plankton Research* **2000**, *22* (1), 91–106. https://doi.org/10.1093/plankt/22.1.91.

(138)   Novarino, G. Nanoplankton Protists from the Western Mediterranean Sea. II. Cryptomonads (Cryptophyceae = Crptomonadea). *Scientia Marina* **2005**, *69* (1), 47–74. https://doi.org/10.3989/scimar.2005.69n147.

(139)   Dhar, B.; Cimarelli, L.; Singh, K.; Brandi, L.; Brandi, A.; Puccinelli, C.; Marcheggiani, S.; Spurio, R. Molecular Detection of a Potentially Toxic Diatom Species. *International Journal of Environmental Research and Public Health* **2015**, *12* (5), 4921–4941. https://doi.org/10.3390/ijerph120504921.

(140)   Tamburri, M. *Protocols for Verifying the Performance of In Situ Chlorophyll Fluorometers*; Protocol ACT PV05-01; Alliance for Coastal Technologies, 2005; p 31.

(141)   *Phytoplankton Manual*; Sournia, A., Ed.; Monographs on oceanographic methodology; Unesco: Paris, 1978.

(142)   Fitoplancton: Metodiche di analisi quali-quantitativa. In *Metodologie di studio del plancton marino*; Socal, G., Buttino, I., Cabrini, M., Mangoni, O., Penna, A., Totti, C., Eds.; 56; ISPRA – Istituto Superiore per la protezione e la ricerca ambientale, 2010; p 658.

(143)   Heil, C. A.; Glibert, P. M.; Fan, C. Prorocentrum Minimum (Pavillard) Schiller. *Harmful Algae* **2005**, *4* (3), 449–470. https://doi.org/10.1016/j.hal.2004.08.003.

(144)   Hashemi, N.; Erickson, J. S.; Golden, J. P.; Jackson, K. M.; Ligler, F. S. Microflow Cytometer for Optical Analysis of Phytoplankton. *Biosensors and Bioelectronics* **2011**, *26* (11), 4263–4269. https://doi.org/10.1016/j.bios.2011.03.042.

(145)   Kheireddine, M.; Ouhssain, M.; Claustre, H.; Uitz, J.; Gentili, B.; Jones, B. H. Assessing Pigment-Based Phytoplankton Community Distributions in the Red Sea. *Frontiers in Marine Science* **2017**, *4*. https://doi.org/10.3389/fmars.2017.00132.

(146)   Barlow, R. G.; Mantoura, R. F. C.; Gough, M. A.; Fileman, T. W. Pigment Signatures of the Phytoplankton Composition in the Northeastern Atlantic during the 1990 Spring Bloom. *Deep Sea Research Part II: Topical Studies in Oceanography* **1993**, *40* (1–2), 459–477. https://doi.org/10.1016/0967-0645(93)90027-K.

(147)   Tamm, M.; Freiberg, R.; Tõnno, I.; Nõges, P.; Nõges, T. Pigment-Based Chemotaxonomy - A Quick Alternative to Determine Algal Assemblages in Large Shallow Eutrophic Lake? *PLOS ONE* **2015**, *10* (3), e0122526. https://doi.org/10.1371/journal.pone.0122526.

(148)   Poryvkina, L.; Babichenko, S.; Leeben, A. Analysis of Phytoplankton Pigments by Excitation Spectra of Fluorescence. In *Proceedings of EARSeL-SIG-Workshop LIDAR Lidar Remote Sensing of Land and Sea*; 1; BIS-Verlag, EARSeL European Association of Remote Sensing Laboratories: Dresden, 2001; Vol. 1, p 348.

(149)   Thrane, J.-E.; Kyle, M.; Striebel, M.; Haande, S.; Grung, M.; Rohrlack, T.; Andersen, T. Spectrophotometric Analysis of Pigments: A Critical Assessment of a High-Throughput Method for Analysis of Algal Pigment Mixtures by Spectral Deconvolution. *PLOS ONE* **2015**, *10* (9), e0137645. https://doi.org/10.1371/journal.pone.0137645.

(150)   Zieger, S.; Mistlberger, G.; Troi, L.; Lang, A.; Confalonieri, F.; Klimant, I. Compact and Low-Cost Fluorescence Based Flow-through Analyzer for Early-Stage Classification of Potentially Toxic Algae and in Situ Semi-Quantification. *Environmental Science & Technology* **2018**. https://doi.org/10.1021/acs.est.8b00578.

(151)   Simis, S. G. H.; Huot, Y.; Babin, M.; Seppälä, J.; Metsamaa, L. Optimization of Variable Fluorescence Measurements of Phytoplankton Communities with Cyanobacteria. *Photosynthesis Research* **2012**, *112* (1), 13–30. https://doi.org/10.1007/s11120-012-9729-6.

(152)   Determann, S.; Lobbes, J. M.; Reuter, R.; Rullkötter, J. Ultraviolet Fluorescence Excitation and Emission Spectroscopy of Marine Algae and Bacteria. *Marine Chemistry* **1998**, *62* (1–2), 137–156. https://doi.org/10.1016/S0304-4203(98)00026-7.

(153)   Seoane, S.; Zapata, M.; Orive, E. Growth Rates and Pigment Patterns of Haptophytes Isolated from Estuarine Waters. *Journal of Sea Research* **2009**, *62* (4), 286–294. https://doi.org/10.1016/j.seares.2009.07.008.

(154)   Jeffrey, S. W. Chlorophyll and Carotenoid Extinction Coefficients. In *Phytoplankton pigments in oceanography: guidelines to modern methods*; International Council of Scientific Unions, Ed.; Monographs on oceanographic methodology; UNESCO Pub: Paris, France, 1997; pp 595–596.

(155)   Callieri, C.; Amicucci, E.; Bertoni, R.; Vörös, L. Fluorometric Characterization of Two Picocyanobacteria Strains from Lakes of Different Underwater Light Quality. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* **1996**, *81* (1), 13–23. https://doi.org/10.1002/iroh.19960810103.

(156)   Mallet, Y.; Coomans, D.; de Vel, O. Recent Developments in Discriminant Analysis on High Dimensional Spectral Data. *Chemometrics and Intelligent Laboratory Systems* **1996**, *35* (2), 157–173. https://doi.org/10.1016/S0169-7439(96)00050-0.

(157)   Casson, A. J.; Luna, E.; Rodriguez-Villegas, E. Performance Metrics for the Accurate Characterisation of Interictal Spike Detection Algorithms. *Journal of Neuroscience Methods* **2009**, *177* (2), 479–487. https://doi.org/10.1016/j.jneumeth.2008.10.010.

(158)   Batista, G. E. A. P. A.; Prati, R. C.; Monard, M. C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *ACM SIGKDD Explorations Newsletter* **2004**, *6* (1), 20. https://doi.org/10.1145/1007730.1007735.

(159) Provost, F.; Fawcett, T.; Kohavi, R. The Case against Accuracy Estimation for Comparing Induction Algorithms. In *Machine learning: proceedings of the fifteenth international conference, Madison, Wisconsin, July 24 - 27, 1998*; Shavlik, J., Ed.; Kaufmann: San Francisco, Calif, 1998; pp 445–453.

(160) Kenneth, R. H. CO-Occurrence of Dinoflagellate Blooms and High PH in Marine Enclosures. *Mar. Ecol. Prog. Ser.* **1992**, No. 86, 181–187. https://doi.org/0171-8630/92/0086/0181/$03.00.

(161) Kana, T. M.; Glibert, P. M. Effect of Irradiances up to 2000 ΜΕ Μ–2 S–1 on Marine Synechococcus WH7803—I. Growth, Pigmentation, and Cell Composition. *Deep Sea Research Part A. Oceanographic Research Papers* **1987**, *34* (4), 479–495. https://doi.org/10.1016/0198-0149(87)90001-X.

(162) Ding, S.; Chen, M.; Gong, M.; Fan, X.; Qin, B.; Xu, H.; Gao, S.; Jin, Z.; Tsang, D. C. W.; Zhang, C. Internal Phosphorus Loading from Sediments Causes Seasonal Nitrogen Limitation for Harmful Algal Blooms. *Science of The Total Environment* **2018**, *625*, 872–884. https://doi.org/10.1016/j.scitotenv.2017.12.348.

(163) Paerl, H.; Joyner, J.; Joyner, A.; Arthur, K.; Paul, V.; O'Neil, J.; Heil, C. Co-Occurrence of Dinoflagellate and Cyanobacterial Harmful Algal Blooms in Southwest Florida Coastal Waters: Dual Nutrient (N and P) Input Controls. *Marine Ecology Progress Series* **2008**, *371*, 143–153. https://doi.org/10.3354/meps07681.

(164) van de Poll, W. H.; Visser, R. J. W.; Buma, A. G. J. Acclimation to a Dynamic Irradiance Regime Changes Excessive Irradiance Sensitivity of *Emiliania Huxleyi* and *Thalassiosira Weissflogii*. *Limnology and Oceanography* **2007**, *52* (4), 1430–1438. https://doi.org/10.4319/lo.2007.52.4.1430.

(165) Latasa, M. Pigment Composition of Heterocapsa Sp. and Thalassiosira Weissflogii Growing in Batch Cultures under Different Irradiances. *Scientia Marina* **1995**, *59* (1), 25–37.

(166) Xia, X.; Liu, H.; Choi, D.; Noh, J. H. Variation of Synechococcus Pigment Genetic Diversity Along Two Turbidity Gradients in the China Seas. *Microbial Ecology* **2018**, *75* (1), 10–21. https://doi.org/10.1007/s00248-017-1021-z.

(167) Morel, A.; Ahn, Y.-H.; Partensky, F.; Vaulot, D.; Claustre, H. <I>Prochlorococcus</I> and <I>Synechococcus</I>: A Comparative Study of Their Optical Properties in Relation to Their Size and Pigmentation. *Journal of Marine Research* **1993**, *51* (3), 617–649. https://doi.org/10.1357/0022240933223963.

(168) Lemaire, E.; Abril, G.; De Wit, R. Distribution of Phytoplankton Pigments in Nine European Estuaries and Implications for an Estuarine Typology. *Download PDF Biogeochemistry* **2002**, *59* (5). https://doi.org/10.1023/A:1015572508179.

13. Curriculum Vitae

Silvia Elisabeth Zieger

Current Address: Steyrergasse 143, 8010 Graz (Austria)

Date of birth: 19.02.1990 in Munich

Nationality: German

E-Mail: silviae.zieger@gmail.com

Profile

My educational and professional background lies in the area of optical sensor development, combining mathematics and analytical chemistry. My field of experience includes the development of analytical instrumentation, their calibration and optimization for various applications in marine science. Besides this, another subject is the modelling of chemical systems with the purpose to enhance our understanding of environmental systems. Moreover, I have experience in organization of scientific events and project management.

Education

03/2016 – present   **PhD candidate** in Natural Science at Graz University of Technology; Topic *"Multiparametric Sensing – from Development to Application in Marine Science"*; Supervision of Prof. I. Klimant

10/2013 – 12/2015   **Master program** in Chemistry at Graz University of Technology**;** The general chemistry Master's Degree focuses on analytical, photochemical and physical analysis and applications. **Thesis**: "*Optical detection module for algae species used as early warning system"*. Supervision of Prof. I. Klimant; Grade*:* 'with distinction'

10/2010 – 09/2013   **Bachelor program** in Environmental System Sciences at the University of Graz and Graz University of Technology (NAWI-project); Environmental System Sciences focuses on analytical and theoretical chemistry. Furthermore, the handling and understanding of complex system dynamics with analytical and mathematical methods is learned. **Chemical thesis**: "*NMR spectroscopic investigations of several naturally occurring monosaccharides"*. Supervision of Prof. Klaus Zangger; Grade*:* 'with distinction'

Mathematical thesis: *"The reasonableness of the nocturnal fall - a mathematical consideration"*; Supervision of Prof. Jens Schwaiger; Grade: 'with distinction'

Professional Training and skills

07/2017	Participation in the 10th Advanced Study Course on Optical Chemical Sensors
06/2016	Organization of the Summer school held in Plentzia, Bilbao (Spain) within the frame of the EU funded SCHeMA project.
06/2013	Participation on the 6th International Seminar in STD for the sustainable technology development

Seminars in 'First aid course', 'Organization of Conferences', 'Mediation Basic', 'Competence training for Teaching Assistants', 'Effective Scientific Writing', 'Data Visualization' and 'Python for Beginners'.

Internships and Collaborations

08/2017 – present	**Technical editor** at Microchimica Acta responsible to the formal reviewing process. As one of several freelancers under the coordination of Prof. O. Wolfbeis, I review submitted manuscripts according to given guidelines.
10/2013 – 10/2017	**EU FP7-Project SCHeMA** – Integrated in Situ Chemical Mapping Probes as part of this EU funded project, I was involved in the development of an optical sensor for the early stage identification of toxin producing algal groups. This project included project meetings around Europe with international partners (both university and industry) and several presentations and reports.
10/2016 – 11/2016	**Internship** at the Institute for Phytoplankton Ecology at the University of the Basque Country (UPV/EHU), Leioa (Spain) supervised by Prof. Sergio Seoane and Aitor Laza. A month-long stay to characterize various algae species stored in their algal cultivation collection.
08/2014 – 09/2014	**Traineeship** at Helmholtz Zentrum München in Munich (Germany); Deepening experiences in capillary electrophoresis (CE) and inductively coupled plasma mass spectroscopy (ICP-MS) for biological samples.
03/2012 – 11/2014	**Research fellowship** at the Institute of Process and Particle Engineering at Graz University of Technology; Project: *»The green heart of Styria«* focusing on the optimization of the process network composition of the Styria model regions for climate and energy under the leadership of Professor M. Narodoslawsky.

07/2012 – 08/2012   **Traineeship** at Analytik Jena AG – Analytical Instrumentation in Jena (Germany)**;** Deepening experiences for instrumental analytics especially for the elementary analysis, TOC and AOX. Handling of several analytical instruments, especially the multi EA4000®.

Awards and Prizes

07/2017   1st Price in the Student Project Competition organized in the frame of the 10th ASCOS summer school on biochemical sensors in medicine held in Trest (Czech Republic).

Publications in peer reviewed journals

Zieger, S., Seoane, S., Laza-Martínez, A., Knaus, A., Mistlberger, G. & Klimant, I.; *Spectral characterization of eight marine phytoplankton phyla and assessing a pigment-based taxonomic discriminant analysis for the in-situ classification of phytoplankton blooms. Environ. Sci. Technol.,* **2018**

Zieger, S., Mistlberger, G., Troi, L., Lang, A., Confalonieri, F., Klimant, I., *Compact and low-cost fluorescence based flow-through analyzer for in situ quantification and early-stage classification of toxic algal blooms*, *Environ. Sci. Technol.,* **2018**

Akter, M.; Auguste, M.; Bir, J.; Blanco, E.; Briaudeau, T.; Dapueto, G.; de Cerio, D.; Coccoli, C.; Creemers, M.; Espino, M.; Gain, D.; Gil-Uriarte, E.; Grimaldi, C.; Khalil, S.M.I.; Lopez, A.; Rementeria, A.; Roch, M.; Rodriguez, A.; Roman, O.; Ruoyu, H.; Abdou, M.; Crespi, M.C.; Botia, M.C.; Fighera, M.; Gil-Diaz, T.; Kowal, J.L.; Luxenburger, F.; Mirasole, Ch.; Pankratova, N.; Penezic, A.; **Zieger, S.;** Tercier-Waeber, M.; Nardin, C.; SCHeMA EU Project Summer School Report (Bilbao June 16 – 17, 2016); CHIMIA International Journal for Chemistry, **2017**, 71(9), 607-610(4)

Tahirbegi, I. B., Ehgartner, J., Sulzer, P., **Zieger, S.,** Kasjanow, A., Paradiso, M., Strobl, M., Bouwes, D., Mayr, T.*; Fast pesticide detection inside microfluidic device with integrated optical pH, oxygen sensors and algal fluorescence. Biosensors and Bioelectronics*, **2017**, *88*, 188-195. doi: 10.1016/j.bios.2016.08.014

Oral presentations

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I.; *A fluorescence based multi-wavelength detection module for toxin producing algae applied in field*; Europt(r)ode XIV - Optical chemical sensors and biosensors Napoli, Italy, March **2018**. (flash presentation)

Zieger, S.E., Mistlberger, G., Troi, L., Lang, A., Holly, C., Klimant, I.; *When the ocean is out of its balance – a fluorescence based monitoring system for toxin producing algae*; DocDays Chemistry 2017, Graz, Austria, **2017**

Zieger, S. E., Mistlberger, G., Troi, L., Lang, A., Holly, C., Klimant, I.; *A Fluorescence Based Miniaturized Detection Module for Toxin Producing Algae*; 2016 AGU Fall Meeting, San Francisco, USA / United States, **2016**

Conference contribution

SCHeMA open and modular in situ sensing solution; Leaflet and contribution to the talk of Marylou Tercier-Waeber at the Oceanology International (Exhibition and Conference), London, Great Britain, 15-17th of March **2016**

Poster presentations

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I*.; A pigment-based taxonomic discrimination analysis for the in-situ classification of (harmful) algal blooms;* XVII Chemometrics in Analytical Chemistry 2018, Halifax, Canada, June **2018**.

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I.; *A fluorescence based multi-wavelength detection module for toxin producing algae applied in field*; DocDays in Graz, Austria, April **2018**.

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I.; *A fluorescence based multi-wavelength detection module for toxin producing algae applied in field*; Europt(r)ode XIV - Optical chemical sensors and biosensors Napoli, Italy, March **2018**.

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I.; *Miniaturised low-cost algae detection module for toxin producing algae*; Europt(r)ode XIII - Optical chemical sensors and biosensors Graz, Austria, March **2016**.

Zieger S. E., Mistlberger G., Troi L., Lang A. and Klimant I.; *Miniaturised low-cost algae detection module for toxin producing algae*; DocDays in Graz, Austria, March **2016**.

Zieger S. E., Mistlberger G., Müller B., Troi L., Holly Ch. and Klimant I*.; Multi-channel miniaturized algae detection unit as an early warning system*.; 13th International Estuarine Biogeochemistry Symposium Bordeaux, France, June **2015**.

References

Prof. Ingo Klimant

Graz University of Technology

Institute of Analytical Chemistry and Food Chemistry

Stremayrgasse 9/II

A – 8010 Graz, Austria

klimant@tugraz.at

+43(0)316 / 873 32500

14. List of Figures

15. List of Tables