ISSN: 2710-0588

Advances in Algal Monitoring Technologies: From Conventional Methods to Fluorescence-Based Intelligent Detection

Shuhan Huang

School of Navigation, Shandong Jiaotong University, Weihai 264209, China

Abstract

Algal monitoring supports water-quality assessment, ecosystem observation, and the timely detection of potentially harmful blooms. However, the technologies used for this purpose differ substantially in analytical target, taxonomic resolution, spatial coverage, and suitability for field deployment. This review critically examines microscopy, extracted-pigment spectrophotometry, chromatography, satellite remote sensing, and optical spectroscopy, with particular emphasis on fluorescence-based monitoring. The technological progression from laboratory chlorophyll determination to continuous in-vivo fluorescence, multi-wavelength excitation, and compact LED-based systems has increased measurement frequency and expanded opportunities for estimating algal group composition. Machine learning further enables classification and quantitative interpretation of spectral and imaging data, although reported laboratory performance does not necessarily transfer to mixed natural communities. Key limitations include spectral overlap, variable cellular pigment content, physiological regulation of fluorescence, interference from suspended particles and dissolved organic matter, calibration differences, and restricted training datasets. Portable systems additionally require demonstrated resistance to fouling and stable operation over extended deployments. Future progress is likely to depend on informative excitation channels, complementary optical measurements, embedded processing, and independent validation across environments. Integrating local sensors with regional remote sensing and targeted laboratory confirmation offers a practical route toward low-cost, real-time, intelligent monitoring while preserving the specificity and traceability required for defensible environmental decisions.

Keywords

Algal monitoring; Phytoplankton; Chlorophyll fluorescence; Fluorescence spectroscopy; LED-induced fluorescence; Remote sensing; Machine learning.

1. Introduction

Phytoplankton monitoring is essential for understanding aquatic ecosystem function and detecting changes that threaten water quality. Microscopic photosynthetic organisms support aquatic food webs, transform inorganic nutrients into organic matter, and contribute substantially to carbon cycling. Marine phytoplankton account for a large share of global primary production, making their distribution relevant to both local ecosystem management and global biogeochemistry [1]. In lakes, reservoirs, rivers, estuaries, and coastal waters, changes in phytoplankton abundance and community composition integrate the effects of nutrient supply, light availability, temperature, mixing, and grazing. Here, algal monitoring refers primarily to planktonic microalgae and cyanobacteria, which are commonly considered together in operational monitoring despite their different evolutionary origins.

Algal blooms arise when population growth and physical accumulation produce elevated biomass. Nutrient enrichment can favor bloom development, while warming, stratification, and

ISSN: 2710-0588

altered hydrology influence which organisms proliferate and when [2, 3]. However, high biomass and harmfulness are not equivalent. Some dense blooms cause ecological damage through shading and oxygen depletion, whereas particular toxin-producing populations can pose risks at comparatively low abundance. Harmful algal blooms can disrupt aquaculture, fisheries, recreation, and drinking-water treatment through toxicity, oxygen stress, and changes in water quality. Consequently, monitoring must distinguish the amount of photosynthetic material from the organisms responsible for a potential hazard. Neither chlorophyll concentration nor an optical bloom signal alone establishes toxin concentration or confirms that a population is harmful.

Interpreting long-term bloom trends also requires consistent observations. A global analysis of reported harmful events emphasized the influence of monitoring effort and changing societal exposure on apparent trends [4]. Satellite observations have independently revealed expansion and intensification of coastal phytoplankton blooms over a defined observation period [5]. These findings address different endpoints and should not be treated as interchangeable evidence of a universal increase in toxic species. Their shared implication is that sampling design, detection thresholds, and the variable being measured must accompany any assessment of bloom occurrence. Monitoring technology therefore shapes both operational decisions and the ecological conclusions drawn from environmental records.

Rapid measurements are needed because blooms can change between scheduled sampling visits, move with currents, and form vertically or horizontally concentrated patches. Laboratory analysis provides valuable taxonomic and chemical information, but delays between collection and reporting may limit its usefulness for an emerging event. A continuously operating probe can reveal short-lived changes, yet samples only a restricted volume and location. Conversely, a satellite image describes a broad surface pattern but cannot provide an uninterrupted record at a water intake or aquaculture enclosure. Effective observing systems combine these complementary scales and connect measurements to a practical response, including targeted sampling and confirmatory analysis [6].

Technology selection is further complicated by the different quantities described as algal abundance. Cell concentration, biovolume, pigment concentration, and carbon biomass are related but not interchangeable. A community containing many small cells can have the same chlorophyll-a concentration as one containing fewer large cells. Pigment content per cell also changes with growth conditions. Likewise, fluorescence intensity depends on both the amount of pigment and its photophysical and physiological state. These distinctions explain why a sensor calibrated against one reference method may disagree with another without either measurement being intrinsically meaningless. The relevant question is whether a measurement is sufficiently specific, timely, and reproducible for its intended use.

This review summarizes the major technologies currently used for algal monitoring, with particular emphasis on the transition from conventional laboratory methods to optical, fluorescence-based, portable, and intelligent monitoring approaches. It is a critical narrative review rather than a systematic meta-analysis. Foundational studies are considered alongside recent work, principally from 2017–2025, on multi-wavelength excitation, compact light sources, spectral interpretation, and field deployment. Methods are evaluated by analytical target, preparation requirements, temporal and spatial coverage, taxonomic resolution, and susceptibility to environmental interference. Particular attention is given to whether advances demonstrated with cultured algae can support reliable measurements in mixed natural communities. The central argument is that compact fluorescence sensors and machine learning are most useful when embedded in a calibrated, complementary monitoring framework.

ISSN: 2710-0588

2. Algal Monitoring Technologies

2.1. Conventional Laboratory Methods

Conventional laboratory methods establish the biological and chemical reference against which rapid sensors are interpreted. Microscopy identifies and counts cells using size, shape, ornamentation, pigmentation, and other morphological features. Depending on the organism and optical preparation, an analyst can resolve major groups, genera, or species, while recognizing colonies, chains, and damaged cells. Such observations provide ecological context that a bulk pigment measurement cannot recover. Nevertheless, taxonomic resolution is not uniform: small cells, morphologically similar taxa, and cryptic species may remain unresolved. Microscopy should therefore be described as a potentially high-resolution reference method, not an infallible species identifier [7, 8].

Counting generally requires a defined sample volume, appropriate preservation or prompt examination, and sufficient observations to represent the population. Sedimentation, concentration, and chamber preparation increase turnaround time, while uneven settling, colony fragmentation, and analyst decisions introduce uncertainty. Rare harmful taxa are especially challenging because the volume examined may contain few or no target cells. Counting more fields improves precision but increases labor. Capital costs can be moderate relative to advanced analytical instrumentation, whereas the recurring cost of skilled personnel is substantial. Automated imaging reduces manual screening and increases throughput, but classification still depends on image quality and reliable expert annotations. It extends microscopy rather than removing the need for taxonomic expertise.

Spectrophotometric chlorophyll measurements replace direct cell observations with an integrated pigment endpoint. Samples are commonly filtered, pigments are extracted into a suitable solvent, and absorbance at selected wavelengths is converted to concentration using solvent-specific equations. The classic spectrophotometric work of Lorenzen established an influential basis for distinguishing chlorophyll and degradation-related pigment contributions [9]. These methods are comparatively accessible, use widely available instruments, and support routine water-quality assessment. Their apparent simplicity, however, applies chiefly to the final absorbance reading. Filtration volume, extraction efficiency, storage conditions, solvent composition, pigment degradation, and background correction all influence the result. Analysis therefore takes longer than the instrument reading alone suggests.

Bulk absorbance offers little direct species discrimination because different taxa share chlorophyll-a and overlapping pigment bands. At low pigment concentrations, the signal may be difficult to distinguish from extraction blanks and instrumental noise; at high concentrations, dilution may be necessary. Extracted pigment concentration also describes the collected sample rather than a continuously observed water volume. For these reasons, spectrophotometry is well suited to laboratory biomass assessment and sensor calibration but is generally unsuitable for unattended, reagent-free in-situ operation. Its reliability depends on validated handling procedures and an uncertainty estimate, rather than on assuming that every extraction provides an exact measure of biomass.

High-performance liquid chromatography (HPLC) separates pigments before detection, providing information that bulk spectrophotometry cannot resolve. Chlorophylls, carotenoids, and selected degradation products can be quantified using retention behavior, absorption spectra, and calibrated standards. Method development by Van Heukelem and Thomas illustrates how separation conditions determine the analytical resolution achievable for complex phytoplankton pigment mixtures [10]. HPLC is consequently valuable for validating optical measurements and investigating community changes. Its disadvantages include higher equipment and consumable costs, solvent handling, specialized maintenance, and the time

ISSN: 2710-0588

required for extraction and chromatographic separation. It remains primarily a laboratory technique even when sampling is automated.

Pigment profiles support chemotaxonomic interpretation, but pigment identity does not map uniquely onto species identity. CHEMTAX estimates group contributions from pigment ratios and an assumed relationship between marker pigments and chlorophyll-a [11]. Environmental acclimation and shared pigments can weaken that relationship. In the Golden Horn Estuary, comparison with microscopy showed that pigment-based estimates were insufficient to characterize the community without microscopic information [8]. Laboratory references also have measurable uncertainty: a European interlaboratory study demonstrated the need to quantify differences in HPLC pigment results before using them to validate other observing systems [12]. The practical role of conventional methods is therefore twofold: providing detailed confirmation and supplying traceable reference measurements. Their slower throughput is justified when the monitoring question requires information that a rapid proxy cannot provide.

2.2. Remote Sensing-Based Monitoring

Satellite remote sensing extends algal monitoring from individual samples to spatially coherent observations of lakes, coastlines, and ocean basins. Ocean color reflects the combined effects of absorption and scattering by water, phytoplankton, non-algal particles, and colored dissolved organic matter (CDOM). Chlorophyll-a retrieval algorithms relate corrected water-leaving optical signals to pigment concentration using empirical, semi-analytical, or data-driven approaches. Established ocean-color band-ratio algorithms demonstrate the utility of consistent spectral relationships across satellite sensors [13]. Their success also illustrates an important restriction: an algorithm performs best where the optical relationships represented in its development data remain applicable.

The Sentinel missions exemplify the trade-off between spatial detail and aquatic spectral information. Sentinel-2 multispectral imagery supports analysis of relatively small water bodies and nearshore gradients, whereas Sentinel-3 Ocean and Land Colour Instrument observations provide spectral sampling designed for water-color applications over broader regions. Pahlevan and colleagues developed a machine-learning approach for chlorophyll-a retrieval from both instruments in inland and coastal waters, addressing the need for comparable products across platforms [14]. Platform choice should therefore reflect water-body dimensions, optical conditions, sampling frequency, and the retrieval objective. A nominally finer image does not automatically provide a more accurate pigment estimate if its spectral information or signal quality is inadequate.

The main advantage of satellite observation is its ability to reveal bloom extent, transport pathways, and spatial heterogeneity that would be difficult to reconstruct from sparse sampling. Repeated imagery supports regional surveillance and retrospective analysis of seasonal and interannual changes. Multi-mission records can extend temporal coverage, provided that differences in sensors and processing are accounted for. The Ocean-Colour Climate Change Initiative demonstrates the importance of harmonization and uncertainty characterization when constructing long-term products [15]. Such records are valuable for identifying where intensified field sampling is warranted. They are less suitable for describing minute-scale fluctuations or subsurface features outside the optically observed surface layer.

Atmospheric correction is a major source of uncertainty because the water-leaving contribution must be separated from atmospheric and surface effects. Clouds, aerosols, glint, land adjacency, and mixed shoreline pixels can reduce data availability or bias retrieved reflectance. The ACIX-Aqua intercomparison showed why atmospheric correction must be evaluated specifically for lakes, rivers, and coastal waters rather than assumed transferable from other environments [16]. Cloud screening also means that repeated satellite coverage is

ISSN: 2710-0588

not equivalent to continuous observation. A bloom may develop or disperse between usable scenes, and time composites may obscure short-lived events. Operational products should therefore report valid observation times and quality flags instead of presenting gap-filled maps as direct measurements.

In optically complex waters, suspended sediment and CDOM can change reflectance independently of algal biomass. Shallow bottoms, floating vegetation, and dense surface accumulations introduce additional complications. Red and near-infrared spectral features can be useful in productive waters, but no single index resolves all combinations of these constituents. Regional calibration or optical water-type selection can improve retrievals, although a model calibrated at one location may retain site-specific dependencies. Field matchups require attention to depth, time, and spatial averaging: a bottle collected at a single point does not necessarily represent the water contributing to a satellite pixel [14, 16].

Species-level interpretation remains more demanding than chlorophyll mapping. Distinctive pigments or scattering properties can support discrimination of some populations, particularly when a bloom dominates the signal. In diverse communities, however, broad spectral similarity limits unique identification. Satellite bloom detection should consequently be interpreted as evidence of a surface optical condition requiring ecological interpretation, rather than as direct confirmation of species or toxicity. Its strongest operational contribution is regional context. Combining that context with local fluorescence records and laboratory confirmation provides a more defensible basis for early warning than relying on any single observation scale.

2.3. Optical and Fluorescence-Based Monitoring

Optical measurements offer a route to rapid assessment because pigments interact with light without requiring their chemical extraction. Absorption, scattering, fluorescence, and imaging describe different aspects of the same assemblage. Absorption provides information about pigment content and spectral composition, whereas scattering is affected by particle size, structure, and refractive properties. Fluorescence measures the re-emission of a fraction of absorbed energy and can be detected against a relatively low optical background. These approaches are complementary, but their outputs remain indirect measures of biological variables. Chlorophyll-a is a useful indicator of photosynthetic material; it cannot by itself establish community composition, cell number, carbon biomass, or toxicity.

The potential for optical group discrimination arises from differences in pigment assemblages. Many green algae contain chlorophyll-b, while numerous diatoms and other chromophyte algae contain chlorophyll-c and characteristic carotenoids. Phycobiliproteins contribute distinctive absorption and fluorescence responses in cyanobacteria and some other groups. These patterns are useful generalizations rather than exclusive taxonomic rules. Accessory pigments alter the wavelengths at which light is absorbed and the transfer of excitation energy toward chlorophyll-a. Consequently, algae sharing a similar red chlorophyll emission band may respond differently when the excitation wavelength changes. The information needed for discrimination can reside in relative excitation responses even when emission peaks overlap [17, 18].

Early fluorometric studies primarily addressed sensitive chlorophyll quantification. Holm-Hansen and colleagues established fluorometric determination as an important analytical alternative to absorbance-based measurements [19]. Lorenzen subsequently demonstrated continuous measurement of in-vivo chlorophyll during shipboard observation, helping move the method from discrete laboratory analysis toward repeated measurements of natural water [20]. The distinction between extracted-pigment fluorometry and in-vivo fluorescence is essential. Extraction changes the molecular environment and removes many cellular influences; intact-cell fluorescence retains the effects of pigment packaging, energy transfer, and

ISSN: 2710-0588

photosynthetic regulation. A calibration suitable for an extract cannot simply be treated as a universal calibration for a living assemblage.

Modern in-vivo fluorometers typically illuminate a water volume and detect emission in a selected spectral band. Fast acquisition, limited preparation, and compatibility with flow-through cells or submersible housings make them attractive for profiling, fixed stations, and mobile surveys. Reagent-free operation can support repeated measurements with little sample consumption. Nevertheless, the detector measures an optical response rather than chlorophyll concentration directly. Conversion requires a calibration linked to appropriate reference measurements. A global assessment of WET Labs ECO sensors by Roesler and colleagues identified substantial calibration-related issues and provided recommendations for improving chlorophyll estimates [21]. Factory calibration should therefore be regarded as an initial instrument characterization, not a guarantee of accuracy across communities and environments.

Fluorescence yield varies because absorbed energy is partitioned among photochemistry, heat dissipation, and fluorescence. Changing illumination can alter non-photochemical quenching and produce diurnal signal variation without an equivalent change in pigment concentration. Nutrient limitation, growth stage, temperature, and community composition also modify the measured response. Measurement protocols must distinguish an assay designed to estimate biomass from an assay designed to assess photosynthetic function. Variable-fluorescence methods provide physiological information through controlled excitation sequences, but their derived parameters are not interchangeable with a simple chlorophyll concentration estimate. Guidance for biogeochemical profiling systems emphasizes calibration, quality control, and environmental interpretation throughout the observing workflow [22].

Multiple excitation channels extend bulk fluorometry by probing the different absorption characteristics of pigment systems. Beutler and colleagues demonstrated differentiation of broad algal populations from in-vivo and in-situ excitation responses [17]. This was an important change in analytical ambition: the objective expanded from estimating total chlorophyll to partitioning its contribution among distinguishable spectral groups. Such groups do not necessarily correspond to individual species or to a complete taxonomic classification. Two species may share similar excitation signatures, while one species can change its signature as its pigment ratios acclimate. Multi-channel results should therefore be described at the resolution supported by the calibration library and independent validation.

The mixture problem is especially important. A first approximation treats the measured response as a weighted combination of reference spectra, permitting estimation of group contributions. This approximation becomes less reliable when self-shading, reabsorption, scattering, or physiological differences alter the responses represented by the library. Proctor and Roesler examined the interpretation of multispectral chlorophyll fluorescence in relation to concentration and composition, underscoring the need to characterize variability within as well as between groups [18]. A good mixture fit does not prove a unique biological solution. Several combinations of spectrally similar populations may reproduce the same measured signal, particularly when only a few excitation channels are available.

Fluorescence spectroscopy retains more information than a single emission-band detector. An emission spectrum describes the response across detection wavelengths at one excitation wavelength; an excitation spectrum varies illumination while monitoring an emission band. An excitation–emission matrix (EEM) combines both dimensions. Its two wavelength axes and intensity values explain the common term three-dimensional fluorescence spectroscopy; it is not three-dimensional spatial imaging. EEMs provide a useful framework for separating pigment-related features from some background contributions and for studying mixture structure. Their practical costs include additional acquisition time, more extensive calibration,

ISSN: 2710-0588

and potentially greater data-processing requirements. Sparse excitation sampling can reduce these costs, but must preserve the features needed for the analytical task.

Recent work has addressed spectral variability through improved reference libraries and computational correction. Wang and colleagues developed an interpolated reference-library approach for rapid analysis of phytoplankton EEMs [23]. A separate study by Wang, Chen, and Wang used deep learning to identify and correct CDOM and turbidity interference in algal fluorescence matrices [24]. These studies illustrate two different strategies: representing biological variability more flexibly and estimating non-algal contributions explicitly. Neither strategy eliminates the need to test mixtures and water matrices outside the development set. In particular, apparent removal of background structure should be checked against independently measured pigments or cell concentrations to ensure that genuine biological variation has not also been removed.

Table 1. Comparison of major algal monitoring technologies

Method	Main principle	Advantages	Limitations	Typical application	In-situ capability	Species discrimination
Microscopy	Cell morphology and counting	Direct counts; detailed morphology	Preparation; slow counting; expert labor	Taxonomic surveys; confirmation	Low; automated imaging differs	High for resolvable taxa
Spectrophotometry	Absorbance of extracted pigments	Mature; accessible instrumentation	Extraction; pigment overlap; limited sensitivity	Laboratory chlorophyll-a measurement	Low	Very low
Chromatography	Separation and quantification of pigments	Multiple pigments; chemical specificity	Extraction; solvents; high cost; batch analysis	Pigment reference data; chemotaxonomy	Low	Indirect group estimates
Remote sensing	Water-color reflectance retrieval	Broad coverage; repeat regional mapping	Clouds; correction errors; surface bias	Bloom extent; regional surveillance	Remote surface observation	Usually low
Fluorescence spectroscopy	Pigment excitation and emission spectra	Rapid; reagent-free; multivariate signals	Physiology; overlap; matrix interference	Frequent biomass and group screening	High with field design	Mainly groups; conditional species
Laser-induced fluorescence	Narrow-band laser excitation	Strong excitation; defined geometry	Optical alignment; limited excitation diversity	Point or stand-off chlorophyll monitoring	High with field design	Limited without added features
LED-induced fluorescence	One or several LED excitation bands	Compact; low power; low source cost	Broad source bands; drift; calibration	Portable and distributed monitoring	High with field design	Groups; depends on channels
ML-assisted spectroscopy	Learned mapping from spectral features	Nonlinear classification and quantification	Training bias; domain shift; uncertain labels	Identification; pigment or cell prediction	Conditional on sensor and model	Only within validated domain

Note: Comparisons are qualitative and application-dependent. In-situ capability refers to suitable deployed configurations, not every implementation. Spectral groups are not equivalent to species. Machine learning (ML) is an analytical layer and may accompany several listed methods. The table synthesizes the evidence discussed in Sections 2 and 3.

ISSN: 2710-0588

Laser-induced fluorescence (LIF) uses a spectrally narrow, directional source that can provide strong excitation and well-defined collection geometry. It supports compact point measurements and, in suitable configurations, remote or stand-off observations. He and colleagues reported a chlorophyll monitoring instrument incorporating ambient-light correction and low-power operating modes [25]. More recent shore-based fluorescence radar work further illustrates the extension of LIF toward online chlorophyll observation [26]. These examples support the feasibility of engineered systems, but their performance remains conditional on geometry, water properties, calibration, and deployment conditions. A narrow excitation line improves source definition; it does not automatically provide enough independent biological information for species identification.

Across fluorescence approaches, natural-water interference remains the central limitation. Turbidity changes excitation and collection paths, CDOM can absorb excitation light and contribute fluorescence, and dense algal suspensions can depart from a linear concentration response. Ambient light increases background and may also change cellular physiology. Cyanobacterial fluorescence probes face further challenges when translating pigment signals into abundance because pigment expression and cell organization vary [27]. Background subtraction, optical filtering, controlled measurement geometry, and supporting turbidity measurements can reduce these effects, but the effectiveness of each correction is conditional. Table 1 compares these optical approaches with laboratory and satellite methods; the categories overlap because a portable fluorescence instrument may use a laser or LEDs and may also incorporate machine learning.

3. Intelligent and Portable Algal Monitoring

3.1. Multi-Wavelength and LED-Induced Fluorescence

Portable monitoring requires an instrument architecture that balances optical information with power, cost, mechanical stability, and maintenance. Broadband lamps coupled to filters or monochromators can provide flexible excitation, but often require more optical and thermal management than a compact field instrument permits. Lasers provide high radiance and narrow emission bands, although available wavelengths, alignment, and source control constrain implementation. LEDs offer a useful alternative because discrete spectral bands can be selected electronically in small, relatively inexpensive assemblies. Their advantages are strongest when the monitoring objective can be addressed with a limited number of informative excitation channels rather than a complete laboratory scan.

The value of multiple LEDs lies in the additional excitation contrast, not simply in increasing illumination intensity. If two populations have similar chlorophyll emission under one excitation band, a second band preferentially absorbed by an accessory pigment may produce a different relative response. Combining channels can therefore help distinguish broad groups and constrain concentration estimates. The SAMSON-related work of Deglint and colleagues combined LED-based spectral measurements with microscopy and deep residual learning, demonstrating the utility of complementary spectral and morphological information [28]. It should be interpreted as evidence for a multimodal imaging approach, not as proof that a few bulk fluorescence channels alone achieve equivalent discrimination.

LED selection must consider the organisms of interest, source bandwidth, detector response, optical filters, and natural-water absorption. There is no universally optimal set of excitation wavelengths. A channel that separates selected laboratory strains may provide little additional information in another community. Furthermore, nominally identical LEDs can differ in spectral output, and source intensity changes with drive conditions and temperature. Sequential excitation requires synchronized acquisition and normalization to the delivered light. Dark measurements, reference standards, and checks for excitation leakage are basic

ISSN: 2710-0588

requirements. Portable spectrometers provide richer emission information, whereas photodiode-based channels can reduce cost and power when their spectral selectivity is adequate.

Simplifying the optics can transfer difficulty to data interpretation. Zhu and colleagues used LED-induced one-dimensional fluorescence spectra with recurrence-plot representations and an improved broad learning system to estimate brown-tide algal cell density [29]. This approach illustrates how compact measurements can be paired with feature transformation. However, reshaping a spectrum into a two-dimensional representation does not create new independent spectral observations. Its value must be established by improved prediction on genuinely independent samples, with comparison against simpler models using the same measured information. Similar caution applies when increasing the number of excitation channels: added channels should improve robustness or identifiability enough to justify acquisition time, component cost, and calibration effort.

Portability and in-situ capability are also distinct. A battery-powered instrument measuring manually collected water is portable, but unattended immersion requires reliable fluid contact, resistance to fouling, environmental sealing, stable reference measurements, and recoverable data storage. Power consumption includes the detector, pumping, communications, and computation as well as illumination. Deployment studies should therefore report total operating requirements and maintenance intervals. Low-power LEDs create an opportunity for distributed monitoring, but the cost of maintaining calibrated observations across many locations may exceed the initial sensor cost. The strongest design is the one that delivers a defined analytical target reliably under its intended field conditions.

3.2. Machine Learning for Algal Identification and Quantification

Machine learning translates multivariate optical measurements into useful outputs when simple peak-intensity relationships are inadequate. Its role depends on the task: classification assigns a taxonomic or spectral label, multi-label classification identifies several possible components, and regression estimates chlorophyll-a or cell concentration. Support vector machines and random forests can use selected spectral features for classification, while partial least squares regression and support vector regression offer alternative mappings for concentration estimation. Convolutional neural networks (CNNs) can exploit neighboring spectral structure or EEM patterns. These are alternative modeling choices rather than a fixed hierarchy in which a deeper model is necessarily better.

Recent fluorescence studies illustrate the range of possible targets. Li and colleagues combined EEM spectroscopy with a multi-label CNN for marine algal identification [30]. Their contribution concerns recognizing constituent classes in single and mixed samples; it should not be broadened into a claim of universally validated species-specific chlorophyll measurement. Chen and colleagues used Zernike-moment features with ensemble learning to predict brown-tide algal concentration [31], while subsequent work explored reconstructed fluorescence representations with CNNs [32]. Together, these approaches show how feature selection and representation can make complex spectral information more tractable. Their reported performance should be interpreted in relation to the organisms, concentration ranges, and validation design actually examined.

The distinction between classification and quantitative transfer is particularly clear in the study by Shi and colleagues [33]. A CNN-based framework combined algal classification with chlorophyll-a determination from fluorescence matrices. For mixed samples, reported mean absolute percentage errors were 6.55–10.56% in ultrapure water, but 21.46–123.37% in a Lake Taihu raw-water background before calibration. Calibration reduced the latter range to 11.86–14.18%. These within-study comparisons provide stronger evidence of matrix sensitivity than a headline classification accuracy alone. They demonstrate why a successful classifier does not

ISSN: 2710-0588

automatically yield transferable concentration predictions and why calibration procedures must be included in any assessment of operational performance.

Validation design can matter more than the choice among plausible algorithms. Repeated spectra from one culture, dilutions of the same stock, or nearby pixels from one image are correlated observations. Randomly distributing them across training and test sets can make a model appear more general than it is. A stronger design separates independent cultures, sampling dates, sites, or instruments before preprocessing and feature selection. Data augmentation should remain within the training partition. Performance reporting should include per-class recall and confusion patterns, along with concentration-dependent error and bias for regression. High correlation alone is insufficient when systematic underestimation near an operational threshold would change the monitoring decision.

Laboratory-to-field transfer introduces organisms, water constituents, and physiological states absent from development data. Small datasets and class imbalance exacerbate the problem, while reference-label uncertainty can limit achievable performance. A field model should therefore allow an unknown or low-confidence outcome instead of assigning every sample to a familiar species. Interpretability tools can help identify whether predictions depend on pigment-related features, background fluorescence, or instrument artifacts, although plausible feature importance does not establish causation. For remote sensing, machine learning similarly requires representative optical conditions and reliable atmospheric correction [14]. The useful objective is validated prediction within a stated domain, accompanied by uncertainty and a procedure for recognizing when that domain has been exceeded.

4. Challenges and Future Perspectives

The first challenge is separating biological change from complex water-matrix effects. In practice, suspended particles, CDOM, bubbles, and illumination history can vary together during a bloom, so correcting one interference independently may be insufficient. Future studies should evaluate designed combinations of these factors and then test the resulting correction on natural samples. Simultaneous absorption, fluorescence, and turbidity measurements can constrain interpretation because they respond differently to pigments and background constituents. This does not require every instrument to become a full laboratory spectrometer. It requires choosing a small set of complementary measurements that resolves the main ambiguity in a specified application.

A second challenge is identifiability in mixed communities. Spectral overlap limits how many independent biological components can be recovered, regardless of computational power. Reference libraries should represent multiple strains, growth stages, and environmental conditions rather than one idealized spectrum per taxon. Future algorithms should report the coarsest reliable group when species-level discrimination is unsupported. Combining fluorescence with absorption and imaging is a promising way to add independent information. The multimodal principle demonstrated by Deglint and colleagues is particularly relevant here [28]. However, additional sensors also introduce alignment, synchronization, and sampling-volume differences, which must be included in the validation rather than ignored during data fusion.

Third, calibration and instrument consistency must become explicit parts of system design. Source aging, wavelength-response differences, optical contamination, and changes in detector gain can all affect comparability. Reference standards can track instrumental stability, while periodically collected natural samples connect the optical record to laboratory pigments and taxonomy. Interlaboratory pigment comparisons show that the reference chain itself has uncertainty [12]. Future sensor evaluations should report repeatability, between-instrument variation, calibration drift, and the consequences of replacing components. A model transferred

ISSN: 2710-0588

between instruments should be tested as a new measurement configuration. Shared protocols will be more useful than nominally identical hardware without documented calibration.

Fourth, model generalization is constrained by the limited diversity and standardization of available datasets. Useful benchmarks should retain raw and processed spectra, excitation settings, instrument response, reference methods, environmental conditions, and biological replication. They should distinguish cultured samples, constructed mixtures, natural-water additions, and unmodified field samples because these represent different levels of deployment realism. Site- and time-separated test sets would make comparisons more informative than isolated random splits. Open, traceable datasets could also support calibration transfer and uncertainty estimation. Where labels are uncertain, group-level annotations or expert confidence scores may be more defensible than forcing a precise species label.

Fifth, long-term stability and the connection to management remain limiting steps. Biofouling changes optical transmission and can create false trends before a sensor fails outright. Antifouling strategies involve trade-offs among cleaning effectiveness, power, material compatibility, and environmental suitability [34]. Field evaluations should include maintenance burden, missing-data behavior, recovery after cleaning, and independent checks over seasonal changes. For distributed networks, a low purchase price is insufficient if frequent manual recalibration or servicing is necessary. Total deployment cost and the fraction of observations meeting quality requirements are more meaningful operational measures than component cost alone.

Future systems are likely to combine low-power multi-wavelength excitation, compact detectors, embedded computation, and selective communication. Edge processing can perform quality checks, detect unusual spectra, and prioritize data transmission when network access is limited. Raw observations and model versions should remain recoverable so that automated decisions can be audited and measurements reprocessed. Adaptive sampling could increase temporal resolution during rapid change, while periodic laboratory sampling maintains taxonomic and chemical specificity. These capabilities should be evaluated through sustained deployments rather than inferred from fast inference on a laboratory computer.

Integration across observation scales offers a practical route to intelligent early warning. Satellite imagery can identify spatially extensive anomalies, fixed probes can resolve local temporal dynamics, and mobile sampling can investigate suspected events. Laboratory identification and, where necessary, toxin analysis can then establish the nature of the risk. Experience with regional harmful-bloom observing systems shows the value of connecting such measurements to stakeholders and response procedures [6]. Future evaluations should measure alert sensitivity, false-alarm frequency, lead time, and data availability under real operating conditions. Portable, low-cost, multi-parameter monitoring will be most credible when its speed is matched by transparent uncertainty and a clear path to confirmation.

5. Conclusion

Algal monitoring technologies address different questions and should be selected according to the variable, scale, and decision of interest. Microscopy remains important when cellular morphology, abundance, and taxonomic interpretation are required. Extracted-pigment spectrophotometry and chromatography provide established quantitative references, with HPLC adding the ability to resolve multiple pigments. Their preparation requirements and reliance on laboratory workflows limit measurement frequency, but their role in confirming observations and maintaining calibration remains essential. Automation should strengthen this reference framework rather than assume it is no longer needed.

Remote sensing provides spatial coverage that point measurements cannot reproduce. It is especially valuable for mapping bloom extent, directing field surveys, and examining regional

ISSN: 2710-0588

change. Its observations are constrained by atmospheric conditions, surface optical complexity, and the mismatch between a satellite pixel and local samples. Fluorescence spectroscopy complements this coverage with rapid, reagent-free measurements that can support frequent or continuous observation. However, fluorescence is influenced by physiology and water properties, and chlorophyll-a remains a biomass-related indicator rather than a complete description of the algal community.

Multi-wavelength excitation increases the information available for distinguishing pigment groups, while LEDs make selected excitation channels compatible with compact and low-power instruments. The achievable discrimination depends on the populations represented in calibration, the independence of the measured features, and the severity of spectral overlap. Machine learning can improve classification and quantitative interpretation when these relationships are nonlinear or multivariate. Its practical contribution must be demonstrated through independent biological samples and testing across water matrices, seasons, and instruments, with explicit treatment of uncertain and unfamiliar observations.

The next stage of development is therefore an integrated monitoring framework linking compact optical sensors, informative excitation channels, intelligent analysis, and reliable reference measurements. Progress should be judged by sustained field performance, transparent uncertainty, maintenance requirements, and the usefulness of observations for timely decisions. Combining portable and in-situ sensing with regional remote sensing and targeted laboratory confirmation offers a balanced route toward real-time monitoring. Such integration can improve coverage and responsiveness while retaining the biological specificity needed to interpret environmental change and assess potential harmful blooms.

References

[1] Field, C. B., Behrenfeld, M. J., Randerson, J. T., et al. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. Science, 281(5374), 237–240. https://doi.org/10.1126/science.281.5374.237.

[2] Huisman, J., Codd, G. A., Paerl, H. W., et al. (2018). Cyanobacterial blooms. Nature Reviews Microbiology, 16(8), 471–483. https://doi.org/10.1038/s41579-018-0040-1

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

[4] Hallegraeff, G. M., Anderson, D. M., Belin, C., et al. (2021). Perceived global increase in algal blooms is attributable to intensified monitoring and emerging bloom impacts. Communications Earth & Environment, 2(1), 117. https://doi.org/10.1038/s43247-021-00178-8

[5] Dai, Y., Yang, S., Zhao, D., et al. (2023). Coastal phytoplankton blooms expand and intensify in the 21st century. Nature, 615(7951), 280–284. https://doi.org/10.1038/s41586-023-05760-y

[6] Anderson, C. R., Berdalet, E., Kudela, R. M., et al. (2019). Scaling up from regional case studies to a global harmful algal bloom observing system. Frontiers in Marine Science, 6, 250. https://doi.org/10.3389/fmars.2019.00250

[7] Lombard, F., Boss, E., Waite, A. M., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. Frontiers in Marine Science, 6, 196. https://doi.org/10.3389/fmars.2019.00196

[8] Dursun, F., Tas, S., & Ediger, D. (2021). Assessment of phytoplankton group composition in the Golden Horn Estuary (Sea of Marmara, Turkey) determined with pigments measured by HPLC-CHEMTAX analyses and microscopy. Journal of the Marine Biological Association of the United Kingdom, 101(4), 649–665. https://doi.org/10.1017/s0025315421000631

[9] Lorenzen, C. J. (1967). Determination of chlorophyll and pheo-pigments: spectrophotometric equations. Limnology and Oceanography, 12(2), 343–346. https://doi.org/10.4319/lo.1967.12.2.0343

ISSN: 2710-0588

[10] Van Heukelem, L., & Thomas, C. S. (2001). Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. Journal of Chromatography A, 910(1), 31–49. https://doi.org/10.1016/s0378-4347(00)00603-4

[11] Mackey, M., Mackey, D., Higgins, H., et al. (1996). CHEMTAX - a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. Marine Ecology Progress Series, 144, 265–283. https://doi.org/10.3354/meps144265

[12] Canuti, E. (2023). Phytoplankton pigment in situ measurements uncertainty evaluation: an HPLC interlaboratory comparison with a European-scale dataset. Frontiers in Marine Science, 10, 1197311. https://doi.org/10.3389/fmars.2023.1197311

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

[14] Pahlevan, N., Smith, B., Schalles, J., et al. (2020). Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. Remote Sensing of Environment, 240, 111604. https://doi.org/10.1016/j.rse.2019.111604

[15] Sathyendranath, S., Brewin, R., Brockmann, C., et al. (2019). An ocean-colour time series for use in climate studies: The experience of the Ocean-Colour Climate Change Initiative (OC-CCI). Sensors, 19(19), 4285. https://doi.org/10.3390/s19194285

[16] Pahlevan, N., Mangin, A., Balasubramanian, S. V., et al. (2021). ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. Remote Sensing of Environment, 258, 112366. https://doi.org/10.1016/j.rse.2021.112366

[17] Beutler, M., Wiltshire, K. H., Meyer, B., et al. (2002). A fluorometric method for the differentiation of algal populations in vivo and in situ. Photosynthesis Research, 72(1), 39–53. https://doi.org/10.1023/a:1016026607048

[18] Proctor, C. W., & Roesler, C. S. (2010). New insights on obtaining phytoplankton concentration and composition from in situ multispectral Chlorophyll fluorescence. Limnology and Oceanography: Methods, 8(12), 695–708. https://doi.org/10.4319/lom.2010.8.0695

[19] Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., et al. (1965). Fluorometric determination of chlorophyll. ICES Journal of Marine Science, 30(1), 3–15. https://doi.org/10.1093/icesjms/30.1.3

[20] Lorenzen, C. J. (1966). A method for the continuous measurement of in vivo chlorophyll concentration. Deep Sea Research and Oceanographic Abstracts, 13(2), 223–227. https://doi.org/10.1016/0011-7471(66)91102-8

[21] Roesler, C., Uitz, J., Claustre, H., et al. (2017). Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. Limnology and Oceanography: Methods, 15(6), 572–585. https://doi.org/10.1002/lom3.10185

[22] Bittig, H. C., Maurer, T. L., Plant, J. N., et al. (2019). A BGC-Argo guide: Planning, deployment, data handling and usage. Frontiers in Marine Science, 6, 502. https://doi.org/10.3389/fmars.2019.00502

[23] Wang, S., Zhang, X., Yin, G., et al. (2025). High-precision and rapid analysis method for three-dimensional fluorescence spectra of phytoplankton based on interpolated reference spectral library. Acta Optica Sinica, 45(10), 1030001. (In Chinese with English abstract). https://doi.org/10.3788/aos250555

[24] Wang, M., Chen, T., & Wang, X. (2023). Rapid correction of turbidity and CDOM interference on three-dimensional fluorescence spectra of live algae based on deep learning. Photonics, 10(6), 627. https://doi.org/10.3390/photonics10060627

[25] He, L., Dai, Y., He, R., et al. (2023). Monitoring instrument of the water chlorophyll-a concentration with laser-induced fluorescent. Transactions of the Chinese Society of Agricultural Engineering, 39(19), 232–240. (In Chinese with English abstract). https://doi.org/10.11975/j.issn.1002-6819.202305163

[26] Li, Y., Fu, Y., Yi, W., et al. (2025). Development of high sensitivity shore-based laser induced fluorescence radar and its application in high-precision online monitoring of chlorophyll

ISSN: 2710-0588

concentration. Optics and Lasers in Engineering, 187, 108883. https://doi.org/10.1016/j.optlaseng.2025.108883

[27] Bertone, E., Burford, M. A., & Hamilton, D. P. (2018). Fluorescence probes for real-time remote cyanobacteria monitoring: A review of challenges and opportunities. Water Research, 141, 152–162. https://doi.org/10.1016/j.watres.2018.05.001

[28] Deglint, J. L., Jin, C., & Wong, A. (2019). Investigating the automatic classification of algae using the spectral and morphological characteristics via deep residual learning. In Image analysis and recognition, ICIAR 2019 (pp. 269–280). Lecture Notes in Computer Science, Vol. 11663. Springer, Cham. https://doi.org/10.1007/978-3-030-27272-2_23

[29] Zhu, Q., Li, X., Liu, J., et al. (2024). Prediction of brown tide algae cell density based on improved 2D-BLS with unthresholded recurrence plots. Acta Optica Sinica, 44(23), 2330003. (In Chinese with English abstract). https://doi.org/10.3788/aos241343

[30] Li, R., Gao, L., Wu, G., et al. (2024). Multiple marine algae identification based on three-dimensional fluorescence spectroscopy and multi-label convolutional neural network. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 311, 123938. https://doi.org/10.1016/j.saa.2024.123938

[31] Chen, Y., Chen, T., Duan, W., et al. (2023). Rapid measurement of brown tide algae using Zernike moments and ensemble learning based on excitation-emission matrix fluorescence. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 294, 122547. https://doi.org/10.1016/j.saa.2023.122547

[32] Chen, Y., Zhang, J., Liu, J., et al. (2024). Predicting brown tide microalgae concentrations using reconstructed fluorescence spectroscopy combined with CNN. Microchemical Journal, 199, 110071. https://doi.org/10.1016/j.microc.2024.110071

[33] Shi, X., Wang, D., Li, L., et al. (2025). Algal classification and chlorophyll-a concentration determination using convolutional neural networks and three-dimensional fluorescence data matrices. Environmental Research, 266, 120500. https://doi.org/10.1016/j.envres.2024.120500

[34] Delgado, A., Briciu-Burghina, C., & Regan, F. (2021). Antifouling strategies for sensors used in water monitoring: Review and future perspectives. Sensors, 21(2), 389. https://doi.org/10.3390/s21020389