Remote Sens. **2015**, *7*, 14781-14805; doi:10.3390/rs71114781

OPEN ACCESS

remote sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Hyperspectral Differentiation of Phytoplankton Taxonomic Groups: A Comparison between Using Remote Sensing Reflectance and Absorption Spectra

Hongyan Xi [1,2,*], Martin Hieronymi [1], Rüdiger Röttgers [1], Hajo Krasemann [1] and Zhongfeng Qiu [2]

[1] Institute for Coastal Research, Helmholtz-Zentrum Geesthacht, Max-Planck-Str. 1, Geesthacht 21502, Germany; E-Mails: martin.hieronymi@hzg.de (M.H.); rroettgers@hzg.de (R.R.); hajo.krasemann@hzg.de (H.K.)

[2] School of Marine Sciences, Nanjing University of Information Science & Technology, 219 Ningliu Road, Nanjing 210044, China; E-Mail: zhongfeng.qiu@nuist.edu.cn

***** Author to whom correspondence should be addressed; E-Mail: hongyan.xi@hzg.de; Tel.: +49-4152-87-1588; Fax: +49-4152-87-1596.

Academic Editors: Saskia Foerster, Véronique Carrere, Michael Rast, Karl Staenz, Deepak R. Mishra and Prasad S. Thenkabail

Received: 29 July 2015 / Accepted: 2 November 2015 / Published: 6 November 2015

Abstract: The emergence of hyperspectral optical satellite sensors for ocean observation provides potential for more detailed information from aquatic ecosystems. The German hyperspectral satellite mission EnMAP (enmap.org) currently in the production phase is supported by a project to explore the capability of using EnMAP data and other future hyperspectral data from space. One task is to identify phytoplankton taxonomic groups. To fulfill this objective, on the basis of laboratory-measured absorption coefficients of phytoplankton cultures ($a_{ph}(\lambda)$) and corresponding simulated remote sensing reflectance spectra ($R_{rs}(\lambda)$), we examined the performance of spectral fourth-derivative analysis and clustering techniques to differentiate six taxonomic groups. We compared different sources of input data, namely $a_{ph}(\lambda)$, $R_{rs}(\lambda)$, and the absorption of water compounds obtained from inversion of the $R_{rs}(\lambda)$) spectra using a quasi-analytical algorithm (QAA). $R_{rs}(\lambda)$ was tested as it can be directly obtained from hyperspectral sensors. The last one was tested as expected influences of the spectral features of pure water absorption on $R_{rs}(\lambda)$ could be avoided after subtracting it from the inverted total absorption. Results showed that derivative analysis of measured $a_{ph}(\lambda)$ spectra performed best with only a few misclassified cultures. Based on $R_{rs}(\lambda)$ spectra, the accuracy of this differentiation decreased but the

performance was partly restored if wavelengths of strong water absorption were excluded and chlorophyll concentrations were higher than 1 mg·m^{-3}. When based on QAA-inverted absorption spectra, the differentiation was less precise due to loss of information at longer wavelengths. This analysis showed that, compared to inverted absorption spectra from restricted inversion models, hyperspectral $R_{rs}(\lambda)$ is potentially suitable input data for the differentiation of phytoplankton taxonomic groups in prospective EnMAP applications, though still a challenge at low algal concentrations.

Keywords: phytoplankton taxonomic groups; EnMAP; remote sensing reflectance; absorption; derivative analysis; QAA

1. Introduction

Optical observations of various water types based on *in situ* measurements and remote sensing data have provided comprehensive information on optical properties and concentrations of optically-significant constituents in aquatic systems. Most notably there have been extensive studies focusing on bio-optical algorithms for estimating the concentration of chlorophyll-*a* as a general proxy for phytoplankton biomass and primary production from water surface reflectance, e.g., [1–3]. Recently, different bio-optical and ecological algorithms have been developed for identifying and differentiating between phytoplankton functional types (PFTs) or size class (PSCs), and taxonomic composition of phytoplankton at the ocean surface, including remote sensing algorithms for monitoring and detecting harmful algal blooms, and for identifying specific phytoplankton species [4–8]. These methods can be summarized into four main types: (1) methods using information on chlorophyll or light absorption to distinguish between PFTs or PSCs [9–11]; (2) spectral response methods based on reflectance anomalies for different PFTs/PSCs (e.g., the PHYSAT approach by Alvain *et al.* [12–14]); (3) absorption-based spectral approaches by deriving a phytoplankton size factor [15,16], through look-up tables [17], by a phytoplankton size discrimination model [18], by the partial least squares regression method [19], or by Differential Optical Absorption Spectroscopy (PhytoDOAS) [20,21]; and (4) a backscatter-based method to infer particle size distribution (PSD) and PSCs [22]. Most approaches mentioned above have been tested globally and applications for using these satellite products have been started. However, validations and adaptions of these approaches to new sensors need to be carried out prior to becoming operational.

With recent advances in optical measurements, comprehensive understanding of the light field within the water, and improvements in satellite sensors, the possibility of taxonomic discrimination of phytoplankton groups has been investigated [23–27]. As satellite sensors expanded from multispectral to hyperspectral detection (e.g., Hyperion, HICO, and the future missions EnMAP [28], PACE, and HyspIRI), the consequently higher number of wavebands, narrower spectral bandwidths and fully-covered range of the visible light spectrum provide more comprehensive remote sensing data on spectral properties of the water reflectance. For instance, as one of the advanced hyperspectral satellite missions, EnMAP is currently in its production phase. One of the major scientific tasks to which EnMAP will contribute is the aquatic ecosystems, not only focusing on oceans but also coastal and

inland waters, regarding various water applications such as improvements on quantification of water constituents and taxonomically identification of algal and phytoplankton groups [28]. The future availability of hyperspectral sensors from space provides a high potential for distinguishing phytoplankton groups by their spectral pigment absorption alone and, thereby, provide a better mapping of the phytoplankton community composition, both for global oceans and regional waters.

Several algal groups have distinct optical properties that are related to their taxonomy and size. The shape of phytoplankton light absorption spectra results from absorption of individual pigments contained in the cell, where the pigment composition is genetically fixed and the concentrations in the cell are influenced by photoacclimation [29]. Due to the genetically fixed pigment composition, taxonomically different algal groups of phytoplankton can be differentiated by analyzing spectral absorption properties. Fourth derivative transformation is typically applied to absorption spectra to enhance spectral features, and then the similarity between the fourth derivative spectra of the targeted phytoplankton and a reference spectrum of a known algal species or taxonomic group is analyzed [30]. Using this method Millie *et al.* successfully detected the potential harmful algae *Gymnodinium breve* (now named *Karenia brevis*) [30] and expanded this application to natural waters [31]. Based on spectral derivative analysis methods, such as principal component analysis (PCA), cluster analysis, and Discriminant Analysis, have been tested for this application [32–34]. Instead of a single parameter, measurements of biomass, pigment composition, and fluorescence excitation were combined with the absorption spectra. These supplementary data have also played an important role in precisely distinguishing phytoplankton groups [25,26,35–38].

Optical remote sensing typically provides water-leaving reflectance, $R_{rs}(\lambda)$ as one measured parameter; this has been used to discriminate phytoplankton communities and to identify some single algae species. Craig *et al.* [4] applied two numerical methods to *in situ* hyperspectral measurements of $R_{rs}(\lambda)$ to assess the feasibility of remote detection of the toxic dinoflagellate, *Karenia brevis*. A quasi-analytical algorithm (QAA) was used to invert $R_{rs}(\lambda)$ to derive phytoplankton absorption $a_{ph}(\lambda)$ [39], then the fourth derivatives of derived $a_{ph}(\lambda)$ were compared to the fourth derivative of a reference *K. brevis* absorption using a similarity index analysis. Similar studies to Craig *et al.* [4] have been carried out subsequently to distinguish phytoplankton types or monitor algal blooms by using multispectral and hyperspectral approaches, band ratios, or empirical relationships between a R_{rs} ratio and typical pigment concentration of specific algae [6,23,40–42]. An ocean reflectance inversion model was also developed for inverting marine inherent optical properties for the use of phytoplankton community structure in order to discriminate *Noctiluca miliaris* and diatoms [43]. Most of the studies on algae detection were proposed to identify only a single species or to differentiate a single phytoplankton group in natural waters.

Due to the diversity of phytoplankton in the global oceanic and coastal waters, it is necessary to spectrally differentiate most common phytoplankton groups using hyperspectral data from advanced optical sensors, such as EnMAP. The aims of the current study are to assess the feasibility of using $R_{rs}(\lambda)$ spectra to differentiate several phytoplankton taxonomic groups, and to compare the performance when using $R_{rs}(\lambda)$ directly and absorption obtained from inversion of the reflectance as input data, so that the suitable input data can be determined. This study is proposed to be a preparation research for an application of phytoplankton group differentiation using EnMAP hyperspectral data.

2. Data and Methods

2.1. Algal Cultures

125 cultures of various algal species from six major phytoplankton taxonomic groups were prepared. These cultures included 19 diatom species (*heterokontophyta* (*bacillariophyceae*)), 13 species of dinophytes (*dinophyta* (*dinophyceae*)), four species of prymnesiophytes (*haptophyta* (*prymnesiophyceae*)), three species of cryptophytes (*cryptophyta* (*cryptophyceae*)), 23 species of chlorophytes (*chlorophyta*), and six species of *cyanobacteria* (*cyanophyceae*)). Additional cultures of three different taxonomic groups that were represented by just a single species were not included in this study, as genetic variability inside a single taxonomic group shall be represented by results of several species inside a group. Diatom species were isolated from water samples taken in the North Sea. Species of the other groups were provided by the Alfred-Wegener-Institute, Helmholtz-Center for Polar and Marine Research, and the Leibnitz-Institute of Freshwater Ecology and Inland Fisheries. The cultures were grown in f/2 medium [44] prepared from filtered North Sea water, in the case of marine species, and modified Waris solution [45] in case of the fresh water species. The algae were grown from single isolated cells in light culturing chambers (Rumed, Germany) at 20 °C under 24 h artificial light (day-light fluorescence tubes) of 50 and 100 µmol photons $m^{-2} \cdot s^{-1}$ photosynthetically available radiation. The different light conditions were chosen to take variations in $a_{ph}(\lambda)$ due to photoacclimation in one species into account. Therefore the original isolate was grown for a few days under 50 µmol photons $m^{-2} \cdot s^{-1}$ until a sufficient cell concentration was reached to divide the culture into two 1L-flasks and then grown for another 5−10 days under the two different light intensities. After this photoacclimation period the cell concentrations were still low and the algae still in the exponential growth phase when sampled. During acclimation and until sampling a good physiological status of the cells was controlled daily by measuring the maximum quantum efficiency of photochemistry with a PhytoPAM (Walz, Germany). Cultures were used only when this efficiency was high and cells can, hence, be considered to be in a healthy state; this is typically the case when cultures are in an exponential growth phase under nutrient replete situation. In case the efficiency was too low for a specific culture, a single cell of that culture was isolated and a new culture established.

2.2. Absorption Measurements and Normalization

The absorption coefficient spectra of phytoplankton, $a_{ph}(\lambda)$ (m^{-1}), were measured with a Point-Source Integration-Cavity Absorption Meter (PSICAM) following the procedures outlined by Röttgers *et al.* [46,47]. Determination of $a_{ph}(\lambda)$ are performed in the spectral range of 350–725 nm with a 2 nm resolution, by measuring the absorption coefficient of the culture sample and subtracting the absorption coefficient of the 0.2 µm-filtrate of the same culture sample. All measurements were done at least in triplicate against pure water as the reference. The PSICAM was calibrated daily against a spectrophotometer (Lambda 800, Perkin-Elmer) using solutions of the colored dye Nigrosine. The PSICAM offers accurate and very sensitive determinations of the absorption coefficient without errors induced by light scattered on the algal cells.

The measured $a_{ph}(\lambda)$ spectra were normalized for further utilization. According to Roesler *et al.* [48], absorption spectra exhibit two kinds of variance: variance in magnitude and variance in spectral shape.

Magnitude variances are due to changes in the spectrally averaged absorption coefficient, A_{ph} (m^{-1}), which is the area of the spectral curve over the pre-defined spectral range and can be expressed as:

$$A_{ph} = \frac{\int_{\lambda_{min}}^{\lambda_{max}} a_{ph}(\lambda)\ (m^{-1})\ d\lambda\ (nm)}{\lambda_{max} - \lambda_{min}(nm)} \quad (1)$$

where λ_{max} and λ_{min} are the integration upper and lower spectral limits. Therefore, for each phytoplankton culture, the absorption spectrum was normalized to the underlying area in the range of 400–700 nm:

$$a_{ph}^{n}(\lambda) = \frac{a_{ph}(\lambda)}{A_{ph}} \quad (2)$$

where $a_{ph}^{n}(\lambda)$ (dimensionless) is the area-normalized absorption curve.

2.3. HydroLight Simulations of Hyperspectral Remote Sensing Reflectance

As one of the most important apparent optical properties, remote sensing reflectance, $R_{rs}(\lambda)$, is commonly used in bio-optical models and ocean color remote sensing for water component retrieval and biomass estimation. In this study the absorption spectra of cultures were measured in the laboratory and based on these absorption spectra, and a radiative transfer model was used to compute radiance distributions through the water column and, finally, $R_{rs}(\lambda)$. The simulations were carried out with HydroLight 5.2 (Sequoia Scientific, Inc., Bellevue, WA, USA) [49,50]. HydroLight allows the user to provide input files that define the inherent optical properties (IOPs) used in a simulation in controlled environments and other environmental parameters,, such as ocean surface wind speed, sun and sky irradiance, sun zenith angle, and so forth.

In the present study, hyperspectral $R_{rs}(\lambda)$ were simulated with the following setups and assumptions taken into account:

- Phytoplankton absorption: eight chlorophyll concentrations (*Chl*) were set for each measured $a_{ph}^{n}(\lambda)$ varying from 0.1 to 100 mg·m^{-3} (0.1, 0.5, 1, 3, 5, 10, 50, and 100 mg·m^{-3}). To determine $a_{ph}(\lambda)$ for these different *Chl* concentrations, we used the empirical relationship by Bricaud *et al.* [51] to calculate $a_{ph}(440)$: $a_{ph}(440) = 0.0654\ [Chl]^{0.728}$. The modeled absorption spectra were obtained by multiplying the normalized absorption spectra with $a_{ph}(440)$, *i.e.*, $a_{ph}^{mod}(\lambda) = a_{ph}(440)\ a_{ph}^{n}(\lambda)$. In the end, eight sets of $a_{ph}^{mod}(\lambda)$ were obtained for these different Chl concentrations.
- Chromophoric dissolved organic matter (CDOM) absorption: the HydroLight default exponential model for CDOM absorption with fixed values at 440 nm, $a_{CDOM}(440)$ (m^{-1}), and a single spectral slope of 0.014 nm^{-1} was used [52]: $a_{CDOM}(\lambda) = a_{CDOM}(440)\exp(-0.014(\lambda - 440))$. In the simulations two CDOM concentrations were used with $a_{CDOM}(440)$ = 0.0 and 0.1 m^{-1} to check how the varying CDOM concentrations do influence the performance of the differentiation.
- Absorption by non-algal particles, $a_{NAP}(\lambda)$ (m^{-1}), was determined using a mass-specific absorption coefficients due to non-algal particles, $a_{NAP}^{*}(\lambda)$ ($m^2 \cdot mg^{-1}$), and a particle mass concentration, which is often referred to as total suspended matter concentrations (*TSM*),

i.e., $a_{NAP}(\lambda) = TSMa^*_{NAP}(\lambda)$. The spectral shape of the a_{NAP}^* used is very similar to the HydroLight standard (average coefficient), but is based on spectrally exponentially-shaped *in situ* measurements from the Baltic Sea and Elbe River (unpublished data). Two *TSM* values were set for different simulation scenarios (*TSM* = 0 and 1 g·m^{-3}).

We used standard settings for these HydroLight simulations [50], with the exception of pure seawater absorption and scattering coefficients, $a_w(\lambda)$ and $b_w(\lambda)$, which were calculated with the Water Optical Properties Processor (WOPP) [53], assuming a temperature of 10 ℃ and salinity of 30 PSU. Plankton and non-algal particles are assumed to scatter like Petzold's average particle scattering (each with a backscatter fraction of 0.0183) [50,54]. Furthermore, the ocean is assumed to be infinitely deep and optically homogeneous. Raman scattering, as well as chlorophyll and CDOM fluorescence are taken into account. Other assumptions regarding the atmosphere included: the sun is in zenith, wind speed is 5 m·s^{-1}, a standard atmosphere with marine aerosols, and a clear sky; this results in an aerosol optical thickness at 550 nm of 0.261. $R_{rs}(\lambda)$ spectra were simulated from 400 nm to 700 nm with 2.5 nm spectral resolution, therefore 4000 $R_{rs}(\lambda)$ spectra were finally obtained for the above scenarios including the different *Chl* and *TSM* concentrations and CDOM absorption coefficients. As done above for $a_{ph}(\lambda)$, area-normalization was applied to each simulated $R_{rs}(\lambda)$ spectrum in the range of 400–700 nm to obtain as set of $R^n_{rs}(\lambda)$ for each $a^n_{ph}(\lambda)$ spectrum.

2.4. Inversion of Absorption Spectra from Simulated $R_{rs}(\lambda)$

As $R_{rs}(\lambda)$ can be obtained directly from satellite sensors, potential applications in phytoplankton groups differentiation using satellite data will rely, firstly, on reflectance data. However there can be two ways to utilize $R_{rs}(\lambda)$ data. The simple one is to apply the differentiation approaches directly to the normalized $R_{rs}(\lambda)$; the other one is to invert absorption spectra from $R_{rs}(\lambda)$ data using bio-optical models and then utilize the inverted absorption spectra for differentiation. In the present study, absorption spectra were inverted from the simulated hyperspectral $R_{rs}(\lambda)$ using the quasi-analytical algorithm (QAA) version 5 as described in Lee *et al.* [39,55,56] for optically-deep waters. Prior to choosing the QAA, other reflectance inversion models had been tested (e.g., semi-analytical algorithms in GIOP model [57]), but showed significant discrepancies between the inverted absorption spectra and measured ones, as in these algorithms the $a_{ph}(\lambda)$ are modeled by derived specific absorption coefficients of phytoplankton or by empirical equations using absorption coefficients at a reference wavelength, which deteriorates the spectra features of pigment composition in the full spectral region. The QAA is simple and quick to apply, as its calculation efficiency is similar to that of empirical models, but its accuracy has been shown to be similar to that of optimization methods [39]. Both absorption coefficient spectra, NAP and CDOM, are characterized by an absorption exponentially decreasing with wavelength without pronounced maxima or minima, thus their absorption will have very little influence on the spectral shape of the fourth derivative spectrum [19]. Compared to that, the inversions using bio-optical models for $a_{ph}(\lambda)$, $a_{CDOM}(\lambda)$ and $a_{NAP}(\lambda)$ usually include uncertainties and errors due to a series of assumptions and empirical relationships between the absorption coefficient and wavelength. The total absorption coefficient is typically a precisely retrieved parameter from QAA inversion [39] and as pure water absorption is relatively accurately known, subtraction of the pure water absorption from total absorption might reduce its deteriorating

effect on the spectral absorption features, though uncertainties introduced by the inversion will somehow influence the quality of inverted absorption data. Therefore, we only inverted the non-water absorption, $a_{pg}(\lambda) = a_{ph}(\lambda) + a_{CDOM}(\lambda) + a_{NAP}(\lambda)$, instead of $a_{ph}(\lambda)$, when using QAA. A detailed description of the mathematical steps involved in the QAA inversion process can be found in [56]. The QAA-inverted $a_{pg}(\lambda)$ spectra were also area-normalized to obtain $a_{pg}^{n}(\lambda)$, as done for the lab-measured $a_{ph}(\lambda)$ and HydroLight-simulated $R_{rs}(\lambda)$.

2.5. Derivative Analysis

Derivative spectroscopy has been widely used in the analysis of hyperspectral data using various computation algorithms [58–60]. It can be applied to hyperspectral measurements of both inherent and apparent optical properties (e.g., absorption spectra, remote sensing reflectance). Derivative analysis enhances spectral features and, thus, better distinguishes subtle features in the spectra. In the present study, a finite divided difference algorithm was used to estimate the derivative spectra by taking the difference of a given spectrum over a sampling interval ($\Delta\lambda$), defined as $\Delta\lambda = \lambda_j - \lambda_i$, where $\lambda_j > \lambda_i$. The first and the *n*th derivative are obtained using equations of $\left.\frac{ds}{d\lambda}\right|_i \approx \frac{s(\lambda_j)-s(\lambda_i)}{\Delta\lambda}$ and $\frac{d^n s}{d\lambda^n} = \frac{d}{d\lambda}\left(\frac{d^{(n-1)}s}{d\lambda^{(n-1)}}\right)$, respectively, where *s* is the spectrum used for the derivative transformation [59]. It is noteworthy that the use of the second derivative and fourth derivative transformation yield different meanings: the second derivative provides qualitative identification of pigments only, whereas the fourth derivative provides quantitative identification [61]. Therefore, the fourth derivative spectra of the absorption are often computed to resolve the positions of the absorption maxima attributable to photosynthetic pigments [30]. To be consistent, the fourth derivative transformation was applied to $a_{ph}^{n}(\lambda)$, $R_{rs}^{n}(\lambda)$ and QAA-inverted $a_{pg}^{n}(\lambda)$ spectra in this study. However, as the derivative computation increases noise in the spectrum, smoothing has to be applied to the data [36,58]. The Savitzky-Golay filter was used for smoothing the original data in which a polynomial order of four and a frame size of 21 were appropriately selected after multiple attempts to determine the best compromise between the noise removal and the ability to resolve the fine spectral information. The Savitzky-Golay filter was selected based on the advantage that the filter exhibits excellent properties of the distribution such as relative maxima and minima. Other smoothing filters by conventional methods, such as a moving average usually distort some spectral features by flattening or shifting [62].

2.6. Similarity Index (SI) and Clustering Analysis

All derivative spectra, those of $a_{ph}^{n}(\lambda)$, simulated $R_{rs}^{n}(\lambda)$, and of QAA-inverted $a_{pg}^{n}(\lambda)$, were compared between the 125 input spectra using a similarity index (SI) analysis as described by Millie *et al.* [30]. In the present study the cosine distance was considered as SI and was computed from the angle between two vectors such that $\mathrm{SI} = \frac{A_1 \cdot A_2}{|A_1| \times |A_2|}$, where A_1 and A_2 are vectors that comprise the two derivative spectra. The SI calculation yields a number from 0 to 1, where 0 indicates no similarity, while 1 indicates absolute similarity between the two spectra.

SI analysis is adequate to differentiate two spectra that may represent, or not, two optically-different spectra/groups, but a single SI number alone is not sufficient for the differentiation of several phytoplankton groups. Therefore, hierarchical cluster analysis (HCA) was used to create a hierarchical

cluster tree and to partition the dataset into clusters using a single linkage (nearest neighborhood) algorithm. The linkage algorithm is based on the cosine distance (*i.e.*, SI) between derivative spectra. HCA traditionally displays a dendrogram to represent the hierarchical tree, with individual observations at one end and the clusters to which the data belong at the other. According to the SI defined here, the closer to one the SI is, the more similar are the features of the two compared spectra. Therefore, spectra with a similar phytoplankton composition are expected to appear closer with larger SI in the cluster tree than those having a very different composition [25]. Cluster trees in the current study were generated by using a free software for scientific data analysis PAST version 2.17 [63].

3. Results

3.1. Derivative Analysis and Clustering of Algal Absorption Spectra

Figure 1a shows some representative $a_{ph}^{n}(\lambda)$ spectra of cultures from the six investigated taxonomic algal groups. These $a_{ph}^{n}(\lambda)$ spectra showed considerable variability in spectral shape from one group to another, indicating significant differences in absorption spectral features. Figure 1b shows fourth derivative transformed spectra of $a_{ph}^{n}(\lambda)$, indicating positions of the absorption maxima attributable to single photosynthetic pigments [35]. The HCA cluster analysis was performed on these fourth derivative spectra of $a_{ph}^{n}(\lambda)$. Note that the result of this clustering is sensitive to the selection of the spectral range, as specific absorption imprints caused by accessory pigments usually occur in a narrower range than 400–700 nm. A sensitivity analysis of spectral regions was performed in detail by Torrecilla *et al.* [25] and here we used a similar procedure to determine the optimal spectral range before performing cluster analysis. The optimal range was determined by choosing low SI between each pair of groups based on the SI calculated for all possible varying ranges from 400 to 700 nm (Figure 2 shows the SI variation between *heterokontophyta* and *cryptophyta* as an example). There were, in total, 15 optimal ranges generated for all pairs of groups from the six phytoplankton groups. Those spectral ranges that also show low SI values, but are too narrow (*i.e.*, close to the 1:1 line as shown in Figure 2), were not considered because they present too little information. As we try to differentiate the six groups simultaneously, due to the fact that for each pair of groups the optimal range may vary, in the present study we combined the 15 spectral ranges to include all the important pigment information. This analysis showed that the spectral range of 430–660 nm gave the best information of all accessory pigments and, thus, was used as the optimal spectral range. Clustering results showed that the fourth derivative spectra of measured $a_{ph}^{n}(\lambda)$ can be used to differentiate the six phytoplankton groups quite precisely. All species of *haptophyta*, *chlorophyta*, *cryptophyta*, and *cyanobacteria* were well grouped together. For the *heterokontophyta*, 27 out of 30 cultures, and for the *dinophyta*, 20 of 21 cultures, were classified together (Figure 3 and Table 1). Using a SI of 0.90, *cryptophyta* could be differentiated; while the SI of 0.95 and 0.96 were the thresholds for *cyanobacteria* and *chlorophyta*, respectively. The highest SI was found between *heterokontophyta* and *dinophyta* (0.98) with a few misclassified cultures in between the two groups.

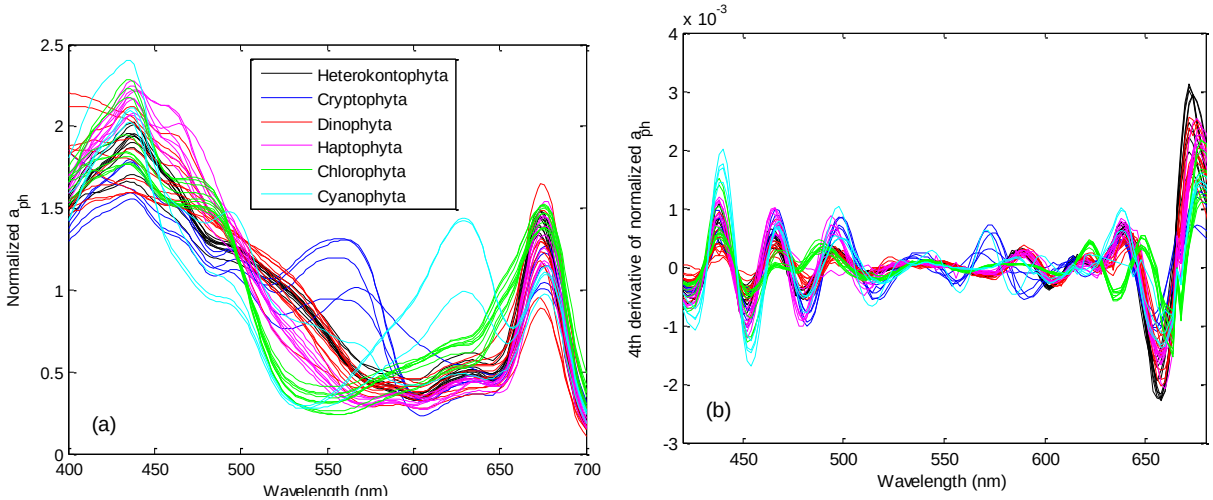

Figure 1. Examples of (**a**) normalized absorption spectra for different phytoplankton groups with (**b**) corresponding fourth-derivative spectra.

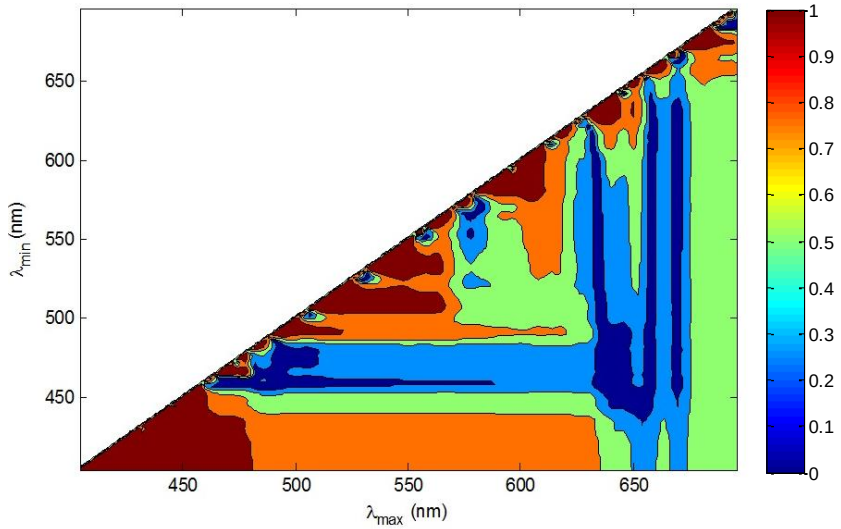

Figure 2. SI variation between *heterokontophyta* and *cryptophyta* for all possible varying ranges from 400 to 700 nm. The optimal spectral range containing most pigment information with low SI is between $\lambda_{min} \approx 430$ nm and $\lambda_{max} \approx 660$ nm.

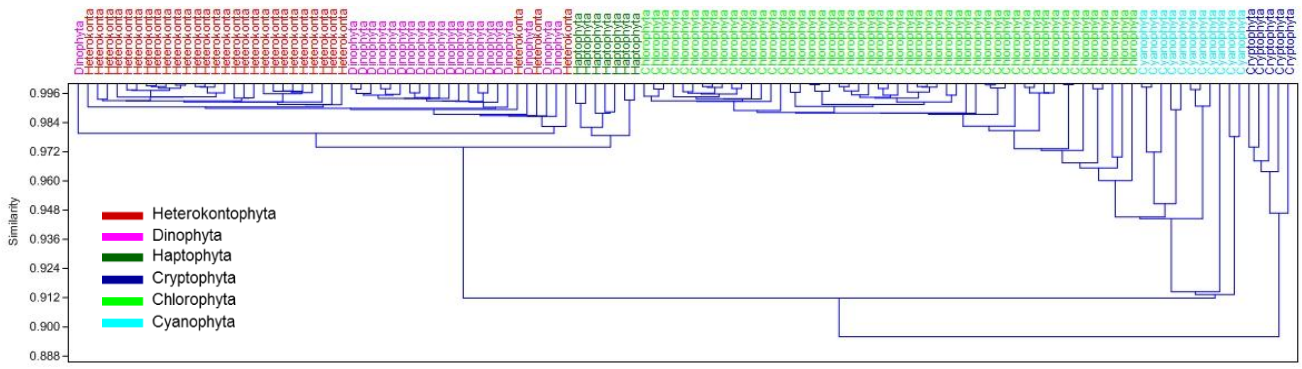

Figure 3. Cluster tree of the six phytoplankton groups generated by using fourth-derivative spectra of measured phytoplankton absorption $a_{ph}(\lambda)$.

Table 1. Clustering accuracy using measured $a_{ph}(\lambda)$, HydroLight-simulated $R_{rs}(\lambda)$, and QAA-inverted $a_{pg}(\lambda)$ ($a_{pg\_QAA}(\lambda)$) spectra for different water types.

	Heterokontophyta	***Dinophyta***	***Haptophyta***	***Cryptophyta***	***Chlorophyta***	***Cyanobacteria***
Measured $a_{ph}(\lambda)$	27/30	20/21	7/7	5/5	51/51	11/11
Simulated $R_{rs}(\lambda)$ (Chl = 0.1 mg·m^{-3})		Mixed		4/5	Mixed	8/11
Simulated $R_{rs}(\lambda)$ (Chl = 0.5 mg·m^{-3})		Mixed		4/5	Mixed	8/11
Simulated $R_{rs}(\lambda)$ (Chl = 1 mg·m^{-3})		Mixed		5/5	51/51	11/11
Simulated $R_{rs}(\lambda)$ (Chl = 5 mg·m^{-3})		Mixed		5/5	51/51	11/11
Simulated $R_{rs}(\lambda)$ (Chl = 10 mg·m^{-3})		Mixed		5/5	51/51	11/11
Simulated $R_{rs}(\lambda)$ (Chl = 50 mg·m^{-3})		Mixed		5/5	51/51	11/11
Simulated $R_{rs}(\lambda)$ (Chl = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, TSM = 1 g·m^{-3})		Mixed		5/5	51/51	11/11
$a_{pg\_QAA}(\lambda)$ (Chl = 1 mg·m^{-3})		Mixed		5/5	49/51	11/11
$a_{pg\_QAA}(\lambda)$ (Chl = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, TSM = 1 g·m^{-3})		Mixed		5/5	49/51	11/11
$a_{pg\_QAA}(\lambda)$ (Chl = 50 mg·m^{-3})	29/30	18/21	Undistinguishable	5/5	51/51	11/11

Numbers "*m*/*n*" indicates that *m* spectra are grouped together from total *n* spectra.

3.2. Derivative Analysis and Clustering on HydroLight-Simulated $R_{rs}(\lambda)$

Figure 4 shows examples of HydroLight-simulated $R_{rs}^{n}(\lambda)$ that are based on individual absorption spectra of the cultures for four *Chl* concentrations varying from 0.1 to 50 mg·m^{-3} and their corresponding fourth-derivative spectra. With increasing *Chl* concentration, the spectral shape of the $R_{rs}^{n}(\lambda)$ and derivative spectra were more distinct between the different taxonomic groups. It is obvious that at low *Chl* concentrations, the $R_{rs}(\lambda)$ were dominated by water absorption features and little variations could be found between $R_{rs}^{n}(\lambda)$ based on absorption spectra of the different cultures. Figure 4 supports this statement and shows that when *Chl* concentration is low (0.1 mg·m^{-3}), little difference between cultures of different taxonomic groups can be seen either in $R_{rs}(\lambda)$ or their derivative spectra, especially in the red to NIR region where absorption by water dominates. The differences between taxonomic groups became more distinct with increasing *Chl* concentration. In order to understand to what extent the *Chl* concentration did influence the differentiation, SI values between different groups were calculated using derivative spectra in the range of 430–620 nm only (to reduce the direct influence of water absorption features). Figure 5 shows an example for the SI

between a single spectrum of each group varying with *Chl* concentration. The SI decreases dramatically with increasing *Chl* concentration between most of the groups, except for the SI between *heterokontophyta*, *dinophyta*, and *haptophyta*, which stayed relatively high compared to the SI between other groups. Figures 4 and 5 indicated that phytoplankton groups differentiation using hyperspectral $R_{rs}(\lambda)$ might not be feasible if *Chl* concentrations are lower than 1 mg·m^{-3} due to high similarity in the derivative spectra.

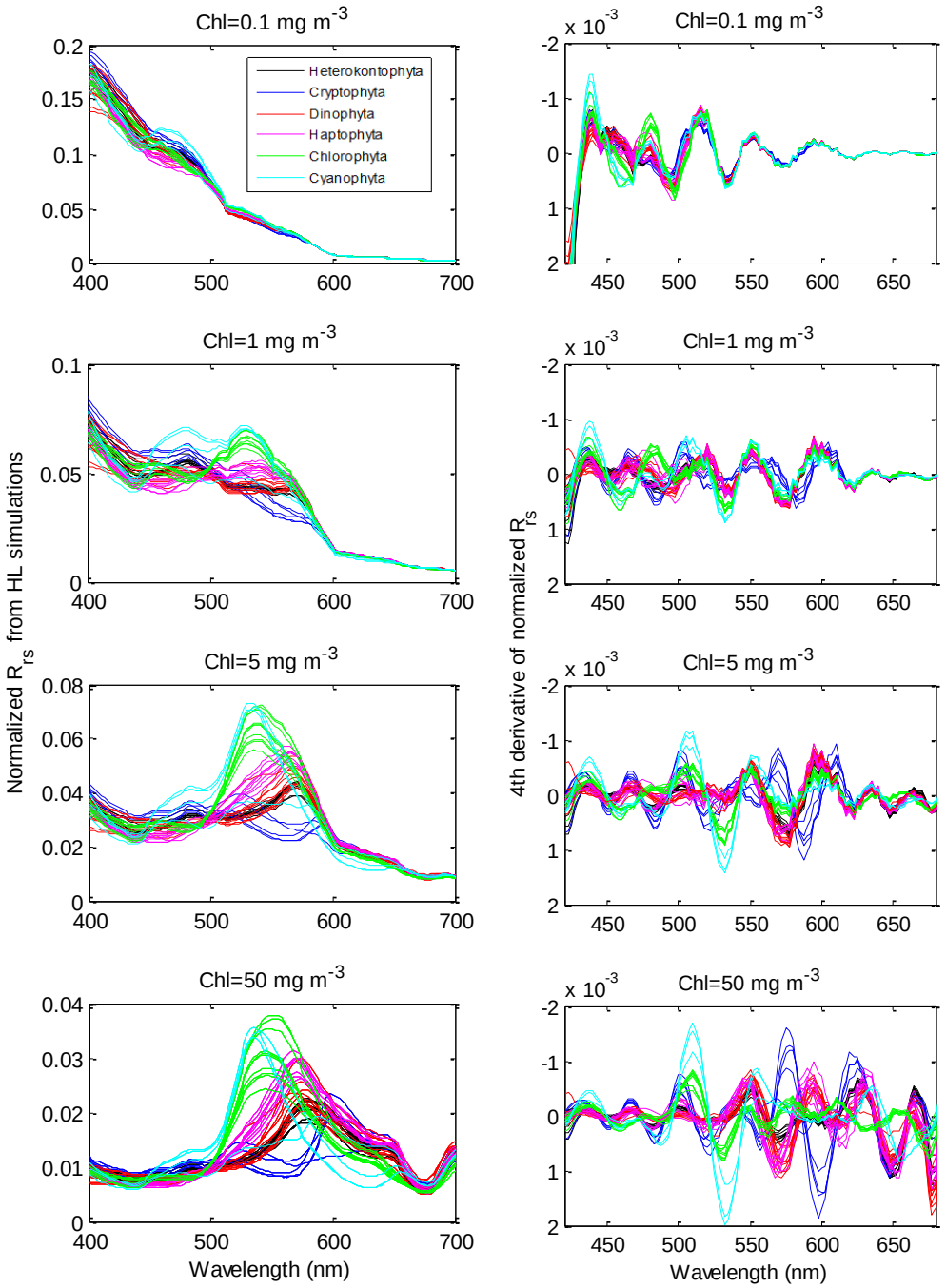

Figure 4. Examples of HydroLight-simulated $R_{rs}^{n}(\lambda)$ with different *Chl* values (**left** panel) and the corresponding fourth-derivative spectra (**right** panel). Note that the y-axis of the derivative spectra is reversely displayed.

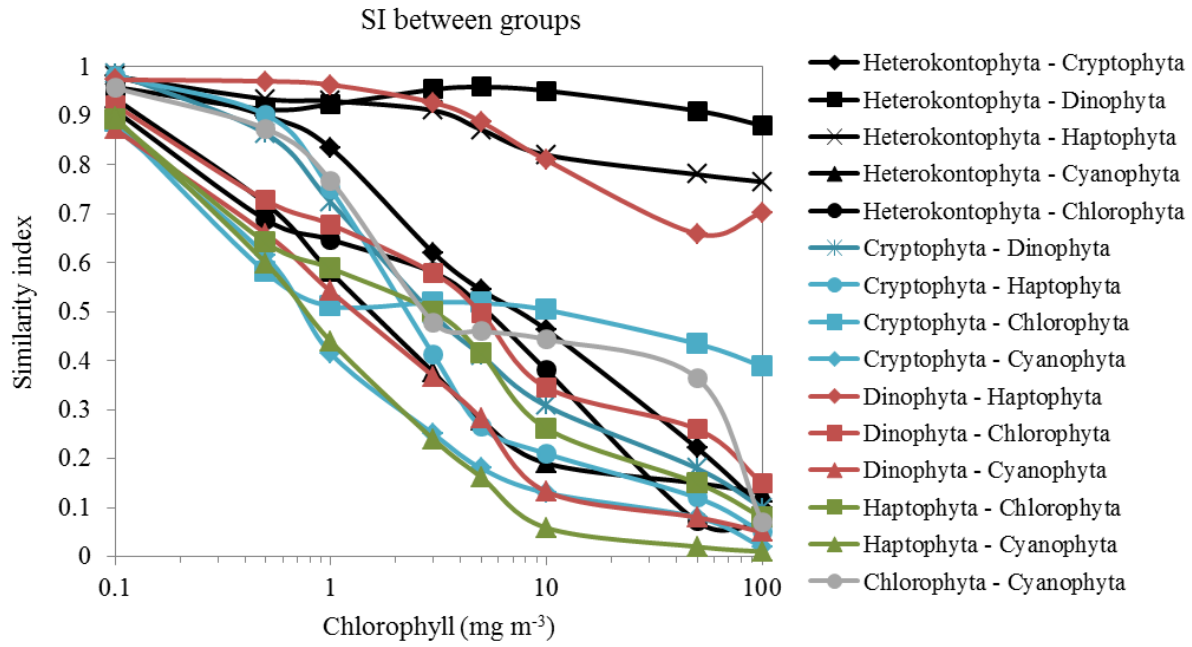

Figure 5. Similarity index between the six taxonomic groups varying with *Chl* concentrations. Note that SI was calculated using the fourth derivative spectra of $R_{rs}^{n}(\lambda)$ within 430–620 nm and between single representative spectra of each group.

To further assess the performance of remote sensing reflectance in differentiating phytoplankton groups, HydroLight-simulated $R_{rs}^{n}(\lambda)$ with a *Chl* concentration of 0.1, 0.5, 1, 5, 10, and 50 mg·m^{-3} were used and the influence by water absorption features at longer wavelengths was reduced by using the fourth derivative spectra in the range of 430–620 nm only. The simulated $R_{rs}(\lambda)$ dataset with *Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, and *TSM* = 1 g·m^{-3} were also tested to assess whether CDOM and *TSM* concentrations influence the differentiation. Cluster trees for different water types are displayed in Figure 6 (cluster trees for water types with *Chl* of 0.5 and 10 mg·m^{-3} were not shown but the accuracy data were listed in Table 1). Results showed that $R_{rs}(\lambda)$ spectra for *Chl* of 0.1 and 0.5 mg·m^{-3} cannot to be used to efficiently differentiate the different taxonomic groups; only *cyanobacteria* and *cryptophyta* were distinct (Table 1 and Figure 6a). When *Chl* = 1 mg·m^{-3} (Figure 6b), the most distinct phytoplankton group are the cryptophytes with the SI of 0.80; the *cyanobacteria* cultures were grouped together at SI = 0.90, but showed larger variation in SI between species in the group due to the above-described variation in spectral absorption. Chlorophytes were well grouped (SI = 0.94), again indicating similar pigment composition among species as presented in the performance of measured $a_{ph}(\lambda)$. When *Chl* is higher than 1 mg·m^{-3}, similar clusters were found but with lower SI thresholds (Figure 6c,d). More subtle differences were shown for *cyanobacteria* cultures, which were separated into two subgroups. When including CDOM and *TSM* in the derivation of $R_{rs}(\lambda)$, the differentiation performance was not visibly deteriorated (Table 1 and Figure 6e). The overall results showed that four main groups can be differentiated using simulated $R_{rs}(\lambda)$ when *Chl* $\geq$ 1 mg·m^{-3}, as all species of *chlorophyta*, *cryptophyta*, and *cyanobacteria* were classified in a single cluster, whereas *heterokontophyta*, *dinophyta*, and *haptophyta* were hardly distinguishable (Table 1 and Figure 6).

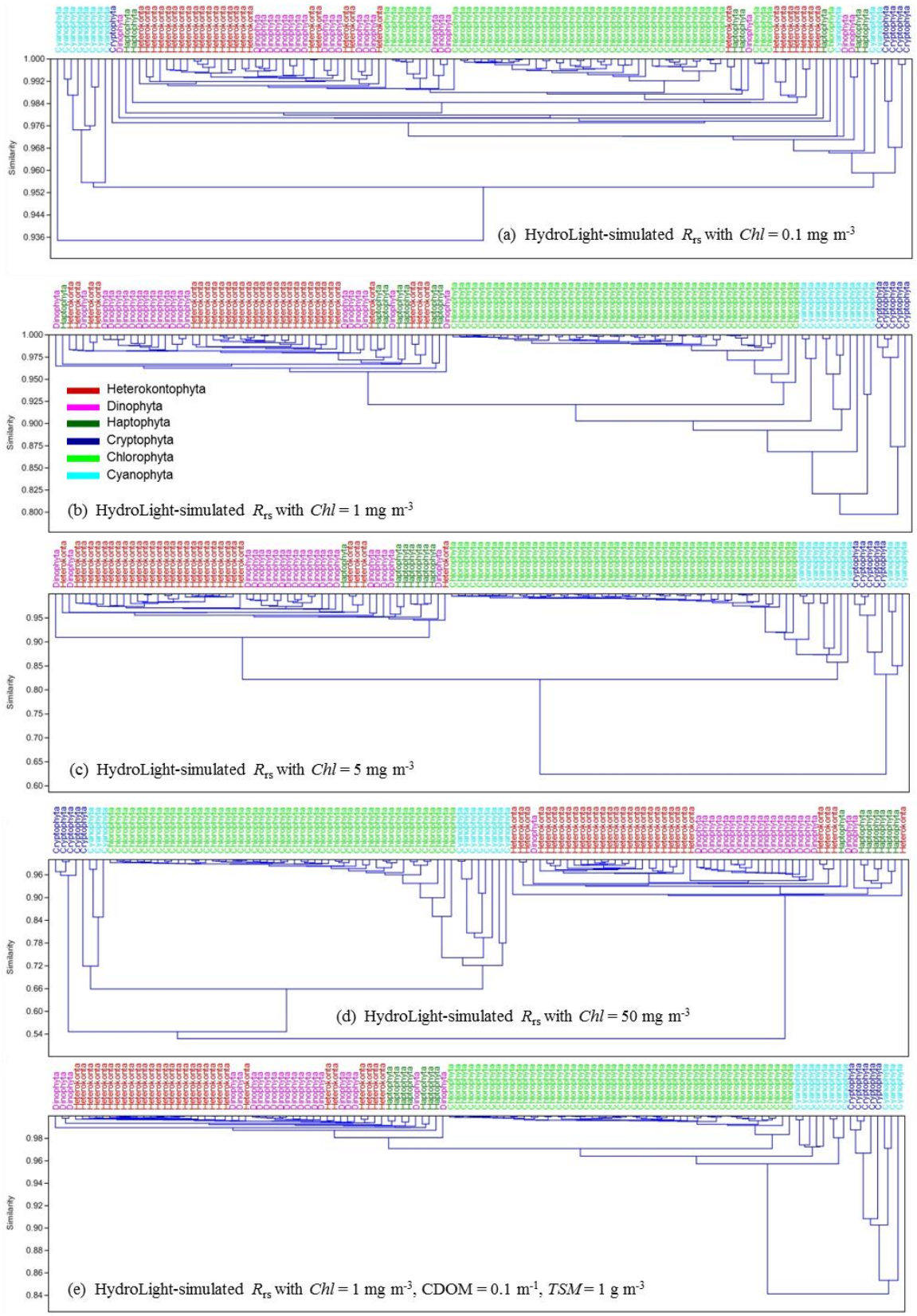

Figure 6. Cluster trees of the six phytoplankton groups generated by using fourth-derivative spectra of HydroLight-simulated $R_{rs}(\lambda)$ for water types with (**a**) *Chl* = 0.1 mg·m^{-3}, (**b**) *Chl* = 1 mg·m^{-3}, (**c**) *Chl* = 5 mg·m^{-3}, (**d**) *Chl* = 50 mg·m^{-3}, and (**e**) *Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, and *TSM* = 1 g·m^{-3}. Color-code legend denotes different phytoplankton taxonomic groups, as shown in (**b**).

3.3. Phytoplankton Group Differentiation Using Absorption Inverted from $R_{rs}(\lambda)$

$R_{rs}(\lambda)$ showed expected weaker differentiation performance compared to that using original $a_{ph}^{n}(\lambda)$ spectra. This is due to the introduction of additional uncertainties during the HydroLight simulation to $R_{rs}(\lambda)$ and due to absorption features by water itself. In this section, the QAA-inverted non-water absorption spectra $a_{pg}(\lambda)$ were analyzed and used to differentiate phytoplankton groups. Its performance was compared with that using $R_{rs}(\lambda)$ data in Section 3.2 to determine the suitable input data for phytoplankton groups differentiation.

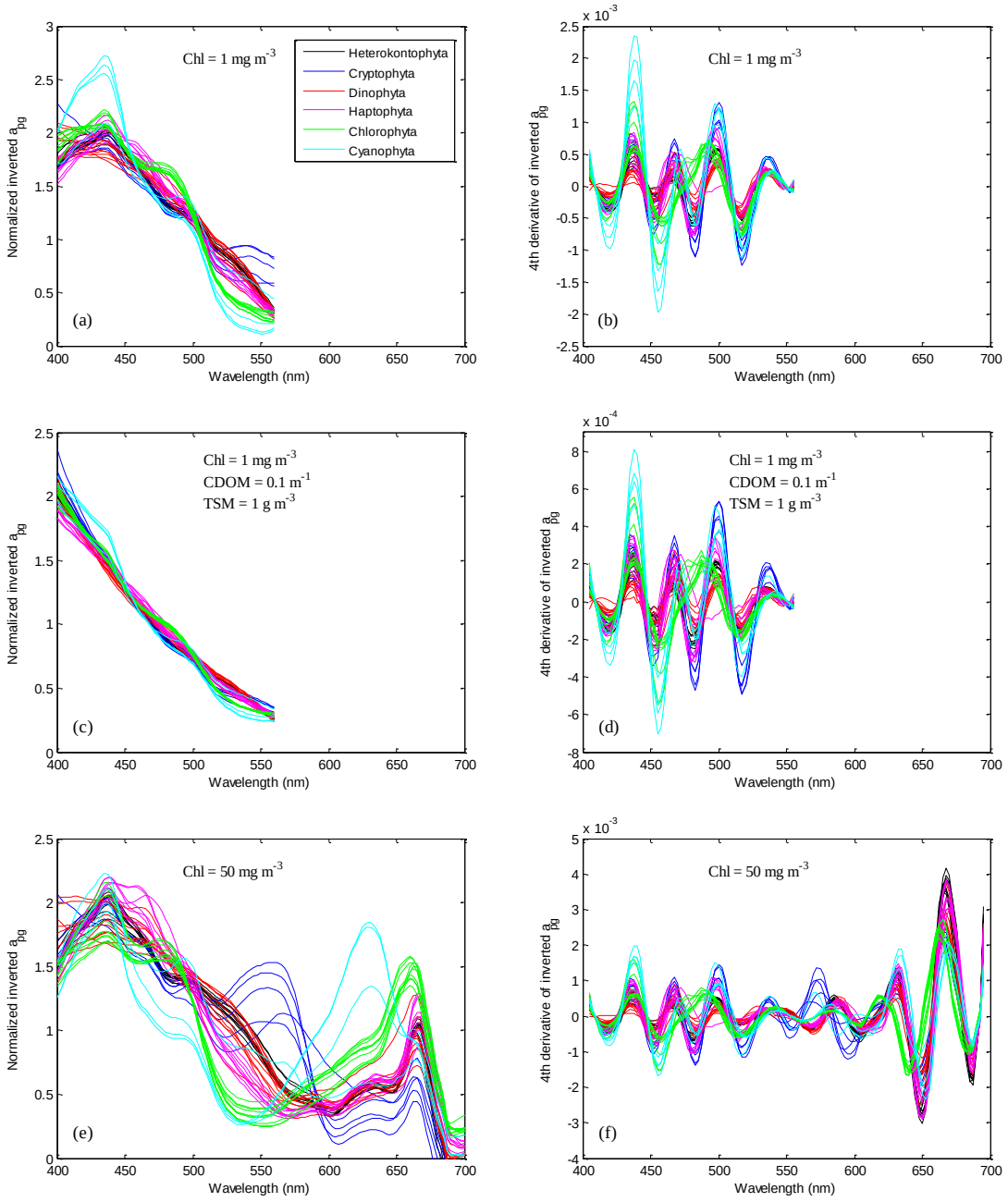

Figure 7. Examples of QAA-inverted absorption spectra of $a_{pg}^{n}(\lambda)$ and the corresponding fourth derivative spectra of the six taxonomic groups for (**a**,**b**) water type I with only *Chl* = 1 mg·m^{-3}, (**c**,**d**) water type II with *Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, and *TSM* = 1 g·m^{-3} for 400–560 nm, and (**e**,**f**) water type III with *Chl* = 50 mg·m^{-3} for 400–700 nm.

In order to verify if the absorption by non-algal particles and CDOM has an influence on the derivative analysis, two different water types were considered to retrieve $a_{pg}(\lambda)$: (I) water with *Chl* = 1 mg·m^{-3}, CDOM and *TSM* were set to zero, and (II) water with *Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, and *TSM* = 1 g·m^{-3}. Note that for water type I and II only data for 400–560 nm were selected because, beyond this range, there is insufficient information to reliably invert $a_{pg}(\lambda)$ from $R_{rs}(\lambda)$ due to errors induced by the strong absorption by water [55]. A water type III was also considered with extremely high *Chl* = 50 mg·m^{-3}, based on the assumption that high absorption by phytoplankton might reduce the QAA inversion errors at longer wavelengths. Results showed that the inverted $a_{pg}(\lambda)$ at longer wavelengths were successfully retrieved for extremely high *Chl* waters (Figure 7e). To assess the performance of the QAA, the measured $a_{pg}(\lambda)$ were compared with the QAA-inverted $a_{pg}(\lambda)$ for the three water types at some discrete wavelengths [55]. Despite of the slight underestimation of the $a_{pg}(\lambda)$ values (slope < 1 in Table 2), the QAA algorithm could give satisfactory retrievals of absorption coefficients from $R_{rs}(\lambda)$ in the considered spectral range (Table 2).

Table 2. Summary of the QAA performance for different water types. A linear regression without the interception term was used between inverted and measured $a_{pg}(\lambda)$, *i.e.*, QAA-inverted $a_{pg}(\lambda)$ = Slope * (measured $a_{pg}(\lambda)$), at four selected bands (410, 440, 490, and 510 nm). The slope, determination coefficient (R^2), root-mean-square error (RMSE), and number of points are shown.

	Slope	**R^2**	**RMSE (m^{-1})**	**N**
Water type I (*Chl* = 1 mg·m^{-3})	0.942	0.961	0.0086	500
Water type II (*Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, *TSM* = 1 g·m^{-3})	0.938	0.948	0.0245	500
Water type III (*Chl* = 50 mg·m^{-3}, an extreme case)	0.953	0.975	0.117	500

Similar to the measured algal absorption spectra, the inverted $a_{pg}(\lambda)$ were normalized and then fourth-derivative-transformed for the cluster analysis. Figure 7a,b shows the normalized $a_{pg}(\lambda)$ ($a_{pg}^{n}(\lambda)$) and the corresponding fourth derivative spectra for water type I and Figure 7c,d shows the same for water type II. In fact, the inverted non-water absorption spectra $a_{pg}(\lambda)$ for water type I correspond to $a_{ph}(\lambda)$ as no absorption by CDOM and non-algal particles was included in the simulation, Nevertheless, the fourth derivative spectra of $a_{pg}^{n}(\lambda)$ for both water types showed little differences in the spectral shape (Figure 7b,d), though their magnitude was different because the inclusion of CDOM and *TSM* for water type II would change the derivative values due to the exponential slopes of CDOM and non-algal particle absorption spectra. As $a_{pg}(\lambda)$ was reasonably inverted at longer wavelengths only for water type III, the normalization and fourth derivative transformation were done within the entire spectral range as shown in Figure 7e,f. Positions of maxima and minima in Figure 7f were in good agreement with that in Figure 7b–d at 420–550 nm and were comparable to that in Figure 1b at 560–660 nm. Cluster analysis was applied to these QAA-inverted $a_{pg}(\lambda)$ data set for water type I and II and a similar clustering was obtained (Figure 8a). The clustering

results showed that four taxonomic groups can be distinguished. All cultures of the *cyanobacteria* and all of the *cryptophyta* were clustered together; 49 of 51 chlorophytes were correctly grouped, however the three other groups (*heterokontophyta*, *dinophyta*, and *haptophyta*,) were mixed within each others' clusters (Table 1 and Figure 8a). Nevertheless, a better clustering was shown for water type III with extremely high *Chl* using the same range of 430–620 nm as done on simulated $R_{rs}(\lambda)$ spectra (Figure 8b). In this case, all cultures of the *chlorophyta*, *cryptophyta*, and *cyanobacteria* were clustered together; only very few cultures of *heterokontophyta* and *dinophyta* were misclassified, but *haptophyta* were undistinguishable.

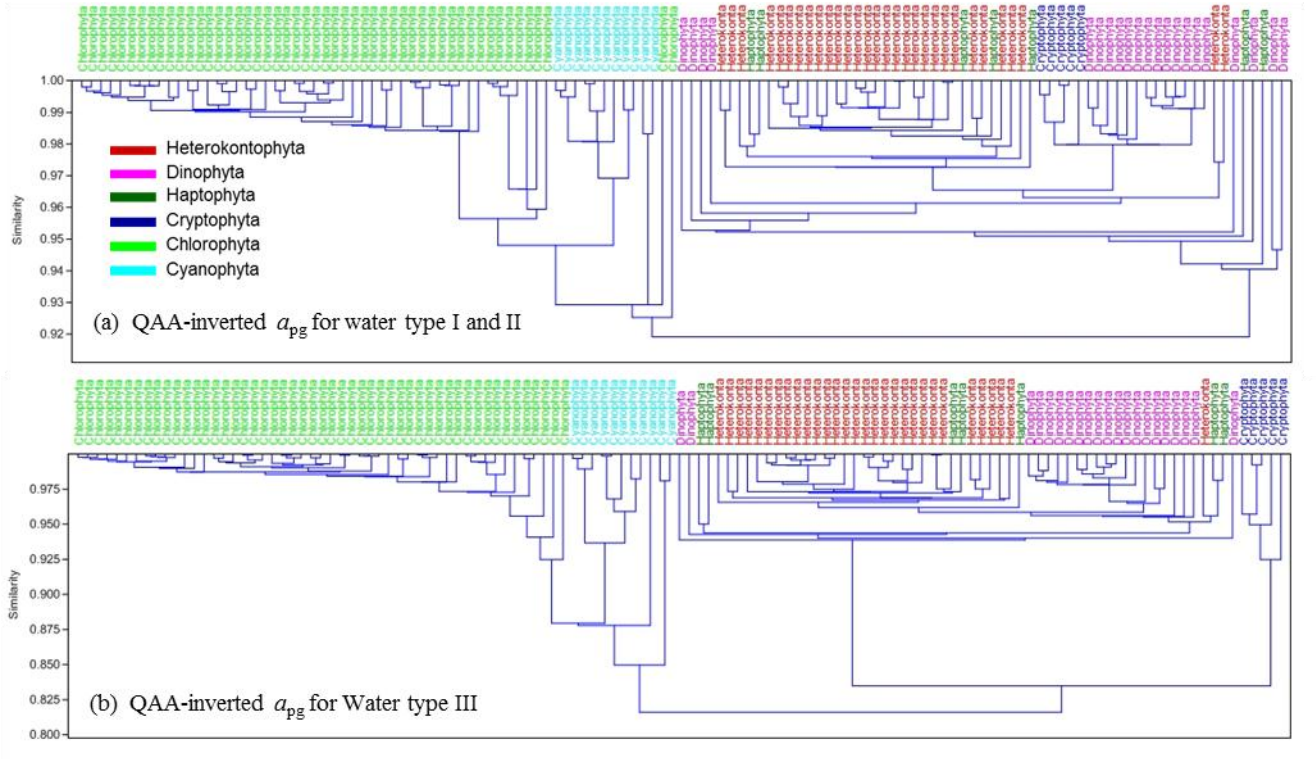

Figure 8. Cluster trees of the six phytoplankton groups generated by using (**a**) fourth derivative spectra of QAA-inverted non-water absorption $a_{pg}(\lambda)$ from simulated $R_{rs}(\lambda)$ for water type I and II, and (**b**) fourth derivative spectra of QAA-inverted non-water absorption $a_{pg}(\lambda)$ from simulated $R_{rs}(\lambda)$ for water type III. Note that the spectral range used in (**a**) is 430–550 nm, and in (**b**) is 430–620 nm.

4. Discussion

4.1. HydroLight Simulations

Part of the scope of this work is to assess how phytoplankton groups are reflected from the hyperspectral remote sensing perspective. As $R_{rs}(\lambda)$ were simulated by HydroLight in this study, two points regarding the HydroLight simulations are discussed in the following. Firstly, as for the scattering properties, besides the fact that scattering properties of hydrosols underlie particular variability and uncertainties [64], some assumptions have been made in the simulations to simplify and unify scattering properties, thus, to strengthen the sensitivity of absorption effects. The volume

scattering function, VSF, exhibits spectral variations and shape changes for different algal species [65]. Depending on the nature and concentrations of planktonic particles in sea water, particle backscatter fractions in the ocean can vary from a fraction of a percent to several percent. If this would be implemented more accurately into the Hydrolight simulations, this would affect the $R_{rs}(\lambda)$ spectra, e.g., [66,67]. However, VSF, particle backscattering, and total scattering of the culture samples have not been measured. Thus, we have no information on the actual particle backscatter fraction b_{bp}/b_p, which is necessary to define the appropriate Fournier-Forand phase function. Various formulas for the backscatter fraction as function of *Chl* can be found in the literature, e.g., [68]. However, it must be remembered that the backscatter fraction correlates poorly with *Chl*, and there can be order-of-magnitude variability in the measured value of b_{bp}/b_p for a given *Chl*. Mobley *et al.* [66] show effects of phase functions on simulation data. They state that the use of a phase function with the correct backscatter fraction could reduce RMS percentage errors in the predicted upwelling irradiance and $R_{rs}(\lambda)$ by roughly an order of magnitude. However, the exact shape of the phase function in backscatter directions does not greatly affect the light field, so long as the overall shape of the phase function does not deviate greatly from the correct shape, and this is provided by using the Petzold average-particle phase function with $b_{bp}/b_p = 0.0183$.

Secondly, with regards to CDOM and non-algal particles absorption, two simple cases were considered, namely, one case without any CDOM and non-algal particles absorption and one with $a_{CDOM}(440) = 0.1\ m^{-1}$ and $TSM = 1\ g{\cdot}m^{-3}$. The CDOM concentration would correspond to roughly 10 $mg{\cdot}m^{-3}$ chlorophyll concentration in the Case-1 parameterization by Morel [69]. According to our routinely-obtained *in situ* dataset mainly from the North Sea, a very high variability of CDOM *versus* *Chl* was found with $a_{CDOM}(440) = 0.1\ m^{-1}$ corresponds to *Chl* between 0.1 and 10 $mg{\cdot}m^{-3}$ (two orders of magnitude). In the open sea, CDOM is mainly a product of phytoplankton degradation. In an extreme and fresh algal bloom event with concentrations of more than 10 $mg{\cdot}m^{-3}$, the used CDOM absorption is probably not unrealistic. Furthermore, CDOM and *TSM* were included in this study to test how they would influence the differentiation performance of the simulated $R_{rs}(\lambda)$, and our results showed that no significant influence was found in results of the cluster analysis (Table 1 and Figure 6).

4.2. Phytoplankton Groups Differentiation Using Absorption and $R_{rs}(\lambda)$ Data—Performance Comparison

The fourth derivative analysis uncovers more distinct the absorption of pigments maxima within the overall absorption spectra [30]. By using the similarity index (SI) with the hierarchical cluster analysis (HCA) it was possible to effectively characterize all absorption spectra and, thus, to allow the detection of differences among phytoplankton taxonomic groups. In the present study, cluster analysis on the fourth derivative of $a_{ph}(\lambda)$ spectra efficiently separated the 125 algal absorption spectra into distinct groups. As expected, the fourth derivative spectra from species of *heterokontophyta* (diatoms) and *dinophyta* were highly similar with SI > 0.975 and few cultures of the two groups were misclassified, due to the known similar pigment composition and absorption spectra of these two groups. Most chlorophyte species showed identical spectral features within the group and, thus, were well-clustered together. The five cryptophyte cultures showed distinct spectral features compared to other groups. The six *cyanobacteria* species (11 cultures) were spectrally rather different from each other, and included green, blue-green, and red colored cultures. For instance, some cultures showed distinct spectral

differences in their absorption, especially at 500–600 nm, induced by several phycobilin pigments (Figure 1). Nevertheless, the current study only focused on the differentiation of phytoplankton taxonomic groups but not yet intended to investigate details at the species level. With significant optical differences to the other groups, *cyanobacteria* are probably easily identifiable using hyperspectral data ([7], and references therein).

For the perspective of applications using data from hyperspectral sensors, HydroLight-simulated $R_{rs}(\lambda)$ for different *Chl* concentrations were tested to differentiate phytoplankton taxonomic groups with this approach. Results revealed that the SI for the fourth derivative of simulated $R_{rs}(\lambda)$ spectra varies largely, as it is highly influenced by the *Chl* concentrations (Figure 5) used in the simulations, leading to an increase of uncertainty in detecting the different taxonomic groups. From the SI variation in Figure 5, a threshold of *Chl* = 1 mg·m^{-3} was primarily determined to more efficiently differentiate the groups when based on $R_{rs}(\lambda)$. The main reason for this less efficient differentiation at low *Chl* concentrations is the dominating influence of spectral features of pure water absorption. Further verification of this threshold was done by using simulated $R_{rs}(\lambda)$ for water types with discrete *Chl* concentrations. Results indicated that four main groups can be differentiated only when *Chl* ≥1 mg·m^{-3} (Table 1). This is also consistent with the SI variation shown in Figure 5, meaning that the derivative analysis and clustering approach using $R_{rs}(\lambda)$ is promising in waters with *Chl* ≥ 1 mg·m^{-3} for differentiating *chlorophyta*, *cryptophyta*, *cyanobacteria*, and *heterokontophyta*/*dinophyta*/*haptophyta* when they are dominating. So far this finding has been noted for the first time by the present study for differentiating multiple phytoplankton taxonomic groups simultaneously; no similar results have been reported in the literature and, thus, cannot be compared. It still remains difficult to distinguish phytoplankton groups with similar optical signatures (e.g., *heterokontophyta* and *dinophyta*) purely depending on reflectance spectra. Approaches to discriminate these two groups during bloom events have been developed using combined data sets of $R_{rs}(\lambda)$, chlorophyll anomaly, absorption, and backscattering spectra from both *in situ* measurements and space, e.g., [42,70–72]. It remains unclear, however, if these techniques are effective in waters beyond their study regions or in non-bloom waters. As an important water components in natural waters, CDOM and *TSM* were also considered in our simulated $R_{rs}(\lambda)$. A test of $R_{rs}(\lambda)$ for water type with *Chl* = 1 mg·m^{-3}, CDOM = 0.1 m^{-1}, and *TSM* = 1 g·m^{-3} showed similar differentiation results with that for water types without CDOM and *TSM* (Figure 6e), indicating that as expected CDOM and *TSM* have an insignificant influence on the differentiation performance also when using simulated $R_{rs}(\lambda)$. The exponential spectral shapes of CDOM and non-algal particles absorption do not influence the derivative analysis.

It is noteworthy that the difficulty in differentiating phytoplankton groups using $R_{rs}(\lambda)$ also has to do with the fundamental difference between the inherent optical properties (IOPs) and apparent optical properties (AOPs), where the latter depend on the ambient light field. According to the discussions above, differentiation of phytoplankton taxonomic groups based on simulated $R_{rs}(\lambda)$ spectra is less effective as that based on $a_{ph}(\lambda)$ spectra. As far as hyperspectral sensors are concerned, one can obtain absorption spectra from $R_{rs}(\lambda)$ by using bio-optical inversion models. However, the differentiation performance of using inverted absorption spectra is much less precise due to uncertainties and errors introduced by the inversion models. The effects seemed to be larger than those induced by the pure water absorption spectral features in $R_{rs}(\lambda)$. Though the QAA-inverted $a_{pg}(\lambda)$ were often in good agreement with the measured ones (Table 2), the QAA showed limitations resulting in the lacking

absorption spectra at longer wavelengths for waters with low *Chl* concentrations due to the influence of the absorption by water itself [4,55]. Compared to simulated $R_{rs}(\lambda)$ for water types I and II, the same approach applied to the QAA-inverted $a_{pg}(\lambda)$ showed slightly poorer differentiation performance, as the QAA does not allow to use longer wavelengths, thus only giving satisfactory inversions of absorption in the range of 400–560 nm when *Chl* is low. This has caused the loss of important pigment information for certain taxonomic groups. For instance, from the measured $a_{ph}^{n}(\lambda)$ and the corresponding fourth-derivative spectra (Figure 1), it was clearly seen that the main spectral difference between *cryptophyta* and other groups was found at 560–600 nm. This deteriorated information at longer wavelengths by QAA resulted in a weak ability to distinguish *cryptophyta* from the three mixed groups (*heterokontophyta*, *dinophyta*, and *haptophyta*) (Table 1 and Figure 8a) and, thus, reduced the differentiation accuracy. The test of an extreme water type III (*Chl* = 50 mg·m^{-3}) suggested that the use of QAA-inverted $a_{pg}(\lambda)$ partly restored the differentiation accuracy when *Chl* is high enough. However, the overall performance of using the inverted absorption spectra was restricted due to the limitations of the inversion model. A recent study in estimating the dominance of diatoms by Isada *et al.* [27] using the derivative spectroscopy approach also suggested that the QAA-inverted absorption spectra is less precise than *in situ* measured absorption. The performance of directly using *in situ* $R_{rs}(\lambda)$ spectra was, however, not assessed in their study.

5. Conclusions and Outlook

In this study we tested the differentiation of phytoplankton taxonomic groups from hyperspectral data by using remote sensing reflectance $R_{rs}(\lambda)$ directly *versus* using absorption spectra derived from $R_{rs}(\lambda)$ by inversion algorithms. This was done to help future implementations of applications for hyperspectral satellite sensors like EnMAP.

When looking at direct differentiation capabilities, the fourth-derivative spectra of measured phytoplankton absorption performed more effectively than that of simulated $R_{rs}(\lambda)$ for six major phytoplankton groups, but the QAA-inverted absorption spectra were less precise than $R_{rs}(\lambda)$. The discrimination of phytoplankton taxonomic groups using $R_{rs}(\lambda)$ data for a spectral analysis is limited by the strong influence of spectral features from pure water absorption. The inversion of $R_{rs}(\lambda)$ to receive pigment absorption and remove water absorption influence is not giving better results due to errors induced by the inversion algorithm. This might change in the future with improved inversion algorithms. Therefore, the use of current hyperspectral remote sensing reflectance directly from hyperspectral sensors for phytoplankton group differentiation is suggested. Additionally, there are difficulties in very low algal concentrations and in discriminating *heterokontophyta*, *dinophyta,* and *haptophyta* due to their similar pigment composition.

The present study is restricted by the use of HydroLight simulations under ideal circumstances with stable CDOM and *TSM* concentrations. Furthermore, it is noteworthy that reflectance spectra of natural waters are far more complex than theoretical simulations, and the quality of the measured spectra are also a matter of the sensor's spectral resolution, the radiometric calibration, and the atmospheric correction, to mention only the most prominent impacts. All of the above issues induce uncertainties and difficulties in identifying phytoplankton groups for natural waters. Future work will include the differentiation capability assessment on phytoplankton group mixtures, the utilization of $R_{rs}(\lambda)$ both

from natural waters and simulated EnMAP hyperspectral images, as well as a more elaborate investigation on the impact of varying water constituent concentrations on $R_{rs}(\lambda)$ spectral features.

Acknowledgments

This work was made possible with the support of the EnMAP project supported by Federal Ministry for Economic Affairs and Energy on the basis of a decision by the German Bundestag. Martin Hieronymi is supported by ESA through a Living Planet Fellowship (project LowSun-OC). The Open Fund at Jiangsu Key Laboratory of Remote Sensing of Ocean Dynamics and Acoustics in NUIST (KHYS1403) is also gratefully acknowledged. We thank Stephen Gehnke for preparing the cultures and measuring the absorption spectra of the phytoplankton cultures. We also thank four reviewers for their valuable comments and suggestions to improve the manuscript.

Author Contributions

Hongyan Xi drafted the manuscript and was responsible for the data analysis and interpretation of the results. Martin Hieronymi simulated the hyperspectral remote sensing reflectance data using HydroLight. Rüdiger Röttgers contributed to the design of the study and revised the draft. Hajo Krasemann provided comments and suggestions to improve the manuscript and contributed to the revision. Zhongfeng Qiu helped with the QAA implementation.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Morel, A. Optical modelling of the upper ocean in relation to its biogenous matter content (case I waters). *J. Geophys. Res.* **1988**, *93*, 10749–10768.
2. O'Reilly, J.E.; Maritorena, S.; Mitchell, B.G.; Siegel, D.A.; Carder, K.L.; Garver, S.A.; Kahru, M.; McClain, C. Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* **1998**, *103*, 24937–24953.
3. Maritorena, S.; Siegel, D.A.; Peterson, A. Optimization of a semi-analytical ocean color model for global scale applications. *Appl. Opt.* **2002**, *41*, 2705–2714.
4. Craig, S.E.; Lohrenz, S.E.; Lee, Z.; Mahoney, K.L.; Kirkpatrick, G.J.; Schofield, O.M.; Steward, R. Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, *Karenia brevis*. *Appl. Opt.* **2006**, *45*, 5414–5425.
5. Kutser, T.; Metsamaa, L.; Strömbeck, N.; Vahtmäe, E. Monitoring *cyanobacteria*l blooms by satellite remote sensing. *Estuar. Coast. Shelf Sci.* **2006**, *67*, 303–312.
6. Astoreca, R.; Rousseau, V.; Ruddick, K.; Knechciak, C.; von Mol, B.; Parent, J.; Lancelot, C. Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters. *J. Plankton Res.* **2009**, *31*, 287–300.

7. Nair, A.; Sathyendranath, S.; Platt, T.; Morales, J.; Stuart, V.; Forget, M-H.E.; Devred, E.; Bouman, H. Remote sensing of phytoplankton functional types. *Remote Sens. Environ.* **2008**, *112*, 3366–3375.
8. Bracher, A.; Hardman-Mountford, N. Overview on algorithms to derive phytoplankton community structure from satellite ocean colour. In Proceedings of the International Ocean Color Science (IOCS) Meeting, Darmstadt, Germany, 6–8 May 2013.
9. Hirata, T.; Aiken, J.; Hardman-Mountford, N.; Smyth, T.J.; Barlow, R.G. An absorption model to determine phytoplankton size classes from satellite ocean color. *Remote Sens. Environ*. **2008**, *112*, 3153–3159.
10. Hirata, T.; Hardman-Mountford, N.J.; Brewin, R.J.W.; Aiken, J.; Barlow, R.; Suzuki, K.; Isada, T.; Howell, E.; Hashioka, T.; Aita-Noguchi, M.; *et al*. Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* **2011**, *8*, 311–327.
11. Brewin, R.J.W.; Sathyendranath, S.; Hirata, T.; Lavender, S.J.; Barciela, R.M.; Hardman-Mountford, N.J. A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Modell.* **2010**, *221*, 1472–1483.
12. Alvain, S.; Moulin, C.; Dandonneau, Y.; Breon, F.M. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. Part I* **2005**, *1*, 1989–2004.
13. Alvain, S.; Moulin, C.; Dandonneau, Y. Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Glob. Biogeochem. Cycles* **2008**, *22*, doi:10.1029/2007GB003154.
14. Alvain, S.; Loisel, H.; Dessailly, D. Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. *Opt. Express* **2012**, *20*, 1070–1083.
15. Ciotti, A.M.; Bricaud, A. Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region of Brazil. *Limnol. Oceanog. Methods* **2006**, *4*, 237–253.
16. Bricaud, A.; Ciotti, A.M.; Gentili, B. Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Glob. Biogeochem. Cycles* **2012**, *26*, GB1010.
17. Mouw, C.B.; Yoder, J.A. Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res. Ocean.* **2010**, *115*, C12018.
18. Fujiwara, A.; Hirawake, T.; Suzuki, K.; Saitoh, S.I. Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences* **2011**, *8*, 3567–3580.
19. Organelli, E.; Bricaud, A.; Antoine, D.; Uitz, J. Multivariate approach for the retrieval of phytoplankton size structure from measured light absorption spectra in the Mediterranean Sea (BOUSSOLE site). *Appl. Opt.* **2013**, *52*, 2257–2273.
20. Bracher, A.; Vountas, M.; Dinter, T.; Burrows, J.P.; Röttgers, R.; Peeken, I. Quantitative observation of *cyanobacteria* and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* **2009**, *6*, 751–764.

21. Sadeghi, A.; Dinter, T.; Vountas, M.; Taylor, B.B.; Altenburg-Soppa, M.; Peeken, I.; Bracher, A. Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci*. **2012**, *8*, 1055–1070.
22. Kostadinov, T.S.; Siegel, D.A.; Maritorena, S. Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* **2009**, *114*, C09015.
23. Lubac, B.; Loisel, H.; Guiselin, N.; Astoreca, R.; Artigas, L.F.; Meriaux, X. Hyperspectral and multispectral ocean color inversions to detect *Phaeocystis globosa* blooms in coastal waters. *J. Geophys. Res.* **2008**, *113*, C06026.
24. Hunter, P.D.; Tyler, A.N.; Présing, M.; Kovács, A.W.; Preston, T. Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution. *Remote Sens. Environ.* **2008**, *112*, 1527–1544.
25. Torrecilla, E.; Stramski, D.; Reynolds, R.A.; Millán-Núñez, E.; Piera, J. Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean. *Remote Sens. Environ.* **2011**, *115*, 2578–2593.
26. Taylor, B.B.; Torrecilla, E.; Bernhardt, A.; Taylor, M.H.; Peeken, I.; Röttgers, R.; Piera, J.; Bracher, A. Bio-optical provinces in the eastern Atlantic Ocean and their biogeographical relevance. *Biogeosciences* **2011**, *8*, 3609–3629.
27. Isada, T.; Hirawake, T.; Kobayashi, T.; Nosake, Y.; Natsuike, M.; Imai, I; Suzuki, K.; Saitoh, S-I. Hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its implications for ocean color remote sensing of diatoms. *Remote Sens. Environ.* **2015**, *159*, 134–151.
28. Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; *et al.* The EnMAP spaceborne imaging spectroscopy mission for Earth Observation. *Remote Sens.* **2015**, *7*, 8830–8857.
29. Hoepffner, N.; Sathyendranath, S. Effect of pigment composition on absorption properties of phytoplankton. *Mar. Ecol. Prog. Ser.* **1991**, *73*, 11–23.
30. Millie, D.F.; Schofield, O.M.; Kirpatrick, 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. *Limnol. Oceanogr.* **1997**, *42*, 1240–1251.
31. Millie, D.F.; Moline, M.A.; Schofield O. Optical discrimination of a phytoplankton species in natural mixed populations. *Limnol. Oceanogr.* **2000**, *45*, 467–471.
32. Johnsen, G.; Samset, O.; Granskog, L.; Sakshaug, E. *In vivo* absorption characteristics in 10 classes of bloom-forming phytoplankton: Taxonomic characteristics and responses to photoadaptation by means of discriminant and HPLC analysis. *Mar. Ecol. Prog. Ser.* **1994**, *105*, 149–157.
33. Moberg, L.; Karlberg, B.; Sorensen, K.; Kallqvist, T. Assessment of phytoplankton class abundance using absorption spectra and chemometrics. *Talanta* **2002**, *56*, 153–160.
34. Hu, X.; Su, R.; Zou, W.; Ren, S.; Wang, H.; Chai, X.; Wang, Y. Research on the discrimination methods of algae based on the fluorescence excitation spectra. *Acta Oceanol. Sin.* **2010**, *29*, 116–128.
35. Smith, C.M.; Alberte, R.S. Characterization of *in vivo* absorption features of *chorophyte*, *phaeophyte* and *rhodophyte* algal species. *Mar. Biol.* **1994**, *118*, 511–521.

36. Roelke, D.L.; Kennedy, C.D.; Weidemann, A.D. Use of discriminant and fourth-derivative analysis with high-resolution absorption spectra for phytoplankton research: Limitations at varied signal-to-noise Ratio and spectral resolution. *Gulf Mex. Sci.* **1999**, *2*, 75–86.
37. Millie, D.F.; Schofield, O.M.E.; Kirpatrick, G.J.; Johnsen, G.; Evens, T.J. Using absorbance and fluorescence spectra to discriminate microalgae. *Eur. J. Phycol.* **2002**, *37*, 313–322.
38. Beutler, M. Spectral Fluorescence of Chlorophyll and Phycobilins as an *in-situ* Tool of Phytoplankton Analysis–Models, Algorithms and Instruments. Ph.D. Thesis, Kiel University, Kiel, Germany, March 2013.
39. Lee, Z.P.; Carder, K.L.; Arnone, R. Deriving inherent optical properties form water color: A multi-band quasianalytical algorithm for optically deep water. *Appl. Opt.* **2002**, *41*, 5755–5772.
40. Von Mol, B.; Ruddick, K.; Astoreca, R.; Park, Y.; Nechad, B. Optical detection of a *Noctiluca scintillans* bloom. *EARSel eProc.* **2007**, *6*, 130–137.
41. Tao, B.; Pan, D.; Mao, Z.; Shen, Y.; Zhu, Q.; Chen, J. Optical detection of *Prorocentrum donghaiense* blooms based on multispectral reflectance. *Acta Oceanol. Sin.* **2013**, *32*, 48–56.
42. Shang, S.; Wu, J.; Huang, B.; Lin, G.; Lee, Z.; Liu, J.; Shang, S. A new approach to discriminate dinoflagellate from diatom blooms from space in the East China Sea. *J. Geophys. Res. Ocean.* **2014**, *119*, 4653–4668.
43. Werdell, P.J.; Roesler, C.S.; Goes, J.I. Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Appl. Opt.* **2014**, *53*, 4833–4849.
44. Guillard, R.R.L. Culture of phytoplankton for feeding marine invertebrates. In *Culture of Marine Invertebrate Animal*; Smith, W.L., Chanley, M.H., Eds.; Plenum Press: New York, NY, USA, 1975; pp. 29–60.
45. McFadden, G.I.; Melkonian, M. Use of Hepes buffer for microalgal culture media and fixation for electron microscopy. *Phycologia* **1986**, *25*, 551–557.
46. Röttgers, R.; Schönfeld, W.; Kipp, P.R.; Doerffer, R. Practical test of a point-source integrating cavity absorption meter: The performance of different collector assemblies. *Appl. Opt.* **2005**, *44*, 5549–5560.
47. Röttgers, R.; Häse, C.; Doerffer, R. Determination of the particulate absorption of microalgae using a point-source integrating-cavity absorption meter: Verification with a photometric technique, improvements for pigment bleaching and correction for chlorophyll fluorescence. *Limnol. Oceanogr. Methods.* **2007**, *5*, 1–12.
48. Roesler, C.S.; Perry, M.J.; Carder, K. Modeling *in situ* phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnol. Oceangr.* **1989**, *34*, 1510–1523.
49. Mobley, C.D. *Light and Water: Radiative Transfer in Natural Waters*; Academic Press: San Diego, CA, USA, 1994.
50. Mobley, C.D.; Sundman, L.K. *HydroLight 5.2—EcoLight 5.2 Users' Guide*; Sequoia Scientific, Inc: Bellevue, WA, USA, 2013.
51. Bricaud, A.; Claustre, H.; Ras, J.; Oubelkheir, K. Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *J. Geophys. Res.* **2004**, *109*, C11010.

52. Bricaud, A.; Morel, A.; Babin, M.; Allali, K.; Claustre, H. Variations of light absorption by suspended particles with chlorophyll-a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *J. Geophys. Res.* **1998**, *103*, 31033–31044.
53. Röttgers, R.; Doerffer, R.; Mckee, D.; Schönfeld, W. The Water Optical Properties Processor. Available online: http://calvalportal.ceos.org/data_access-tools (accessed on 18 February 2015).
54. Petzold, T.J. *Volume Scattering Functions for Selected Ocean Waters*; SIO Ref. 72–78; Scripps Institution of Oceanography: La Jolla, CA, USA, 1972.
55. Lee, Z.P.; Carder, K.L. Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. *Remote Sens. Environ.* **2004**, *89*, 361–368.
56. Lee, Z.P.; Lubac, B.; Werdell, J.; Arnone, R. An Update of the Quasi-Analytical Algorithm (QAA_v5). Available online: http://www.ioccg.org/groups/software.html (accessed on 30 July 2014).
57. Werdell, J.P.; Franz, B.A.; Bailey, S.W.; Feldman, G.C.; Boss, E.; Brando, V.E.; Dowell, M.; Hirata, T.; Lavender, S.J.; Lee, Z.P.; *et al*. A generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* **2013**, *52*, 2019–2037.
58. Butler, W.L; Hopkins, D.W. An analysis of fourth derivative spectra. *Photochem. Photobiol.* **1970**, *12*, 451–456.
59. Tsai, F.; Philpot, W.D. Derivative analysis of hyperspectral data. *Remote Sens. Environ.* **1998**, *66*, 41–51.
60. Torrecilla, E.; Piera, J.; Vilaseca, M. Derivative analysis of hyperspectral oceanographic data. In *Advances in Geoscience and Remote Sensing*; Jedlovec, G., Ed.; InTech: Rijeka, Croatia, 2009; pp. 597–618.
61. Bidigare, R.R.; Morrow, J.H.; Kiefer, D.A. Derivative analysis of spectral absorption by photosynthetic pigments in the western Sargasso Sea. *J. Mar. Res.* **1989**, *47*, 323–341.
62. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639.
63. Hammer, Ø.; Harper, D.A.T.; Ryan, P.D. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Available online: http://palaeo-electronica.org/2001_1/past/issue1_01.htm (accessed on 10 February 2014).
64. McKee, D.; Chami, M.; Brown, I.; Calzado, V.S.; Doxaran, D.; Cunningham, A. Role of measurement uncertainties in observed variability in the spectral backscattering ratio: A case study in mineral-rich coastal waters. *Appl. Opt.* **2009**, *48*, 4663–4675.
65. Tan, H.; Doerffer, R.; Oishi, T.; Tanaka, A. A new approach to measure the volume scattering function. *Opt. Express* **2013**, *21*, 18697–18711.
66. Mobley, C.D.; Sundman, L.K.; Boss, E. Phase function effects on oceanic light fields. *Appl. Opt.* **2002**, *4*, 1035–1050.
67. Dierssen, H.M.; Kudela, R.M.; Ryan, J.P.; Zimmerman, R.C. Red and black tides: Quantitative analysis of water—Leaving radiance and perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments. *Limnol. Oceanogr.* **2006**, *51*, 2646–2659.
68. Twardowski, M.S.; Boss, E.; MacDonald, J.B.; Pegau, W.S.; Barnard, A.H.; Zaneveld, J.R.V. A model for estimating bulk refractive index from optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *J. Geophys. Res.* **2001**, *106*, 14129–14142.

69. Morel, A. Are the empirical relationships describing the bio-optical properties of case 1 waters consistent and internally compatible? *J. Geophys. Res. Ocean.* **2009**, *114*, C01016.
70. Tomlinson, M.C.; Stumpf, R.P.; Ransibrahmanakul, V.; Truby, E.W.; Kirkpatrick, G.J.; Pederson, B.A.; Vargo, G.A.; Heil, C.A. Evaluation of the use of SeaWiFS imagery for detecting *Karenia brevis* harmful algal blooms in the eastern Gulf of Mexico. *Remote Sens. Environ.* **2004**, *91*, 293–303.
71. Tomlinson, M.C., Wynne, T.T.; Stumpf, R.P. An evolution of remote sensing techniques for enhanced detection of the toxic dinoflagellate, *Karenia brevis*. *Remote Sens. Environ*. **2009**, *113*, 598–609.
72. Cannizzaro, J.P.; Carder, K.L.; Chen, F.R.; Heil, C.A.; Vargo, G.A. A novel technique for detection of the toxic dinoflagellate, *Karenia brevies*, in the Gulf of Mexico from remotely sensed ocean color data. *Cont. Shelf Res.* **2008**, *28*, 137–158.