remote sensing

MDPI

Article

Improving the Remote Sensing Retrieval of Phytoplankton Functional Types (PFT) Using Empirical Orthogonal Functions: A Case Study in a Coastal Upwelling Region

Marco Correa-Ramirez [1,2,*], Carmen E. Morales [3,4], Ricardo Letelier [5], Valeria Anabalón [4,6] and Samuel Hormazabal [2,4]

1 Programa de Geociencias, Instituto de Investigaciones Marinas y Costeras INVEMAR, Santa Marta 470006, Colombia
2 Escuela de Ciencias del Mar, Pontificia Universidad Católica de Valparaiso, Valparaiso 2340000, Chile; samuel.hormazabal@pucv.cl
3 Departamento de Oceanografía, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción 4070386, Chile; camorale@udec.cl
4 Instituto Milenio de Oceanografía (IMO-Chile), Concepción 4030000, Chile; vanabalo@gmail.com
5 College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA; letelier@coas.oregonstate.edu
6 Instituto de Oceanografía y Cambio Global, Universidad de Las Palmas de Gran Canaria, Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, Spain
* Correspondence: marco.correa@invemar.org.co; Tel.: +57-301-7400508

Received: 12 January 2018; Accepted: 14 March 2018; Published: 22 March 2018

Abstract: An approach that improves the spectral-based PHYSAT method for identifying phytoplankton functional types (PFT) in satellite ocean-color imagery is developed and applied to one study case. This new approach, called PHYSTWO, relies on the assumption that the dominant effect of chlorophyll-a (Chl-a) in the normalized water-leaving radiance (nLw) spectrum can be effectively isolated from the signal of accessory pigment biomarkers of different PFT by using Empirical Orthogonal Function (EOF) decomposition. PHYSTWO operates in the dimensionless plane composed by the first two EOF modes generated through the decomposition of a space–nLw matrix at seven wavelengths (412, 443, 469, 488, 531, 547, and 555 nm). PFT determination is performed using orthogonal models derived from the acceptable ranges of anomalies proposed by PHYSAT but adjusted with the available regional and global data. In applying PHYSTWO to study phytoplankton community structures in the coastal upwelling system off central Chile, we find that this method increases the accuracy of PFT identification, extends the application of this tool to waters with high Chl-a concentration, and significantly decreases (~60%) the undetermined retrievals when compared with PHYSAT. The improved accuracy of PHYSTWO and its applicability for the identification of new PFT are discussed.

Keywords: phytoplankton functional types (PFT); ocean color; modified PHYSAT method; empirical orthogonal functions (EOF); coastal upwelling waters

1. Introduction

Phytoplankton Functional Types (PFTs) represent an operational division that combines highly diverse taxa (species level) into groups which share traits (morphological, physiological, behavioral, and/or life-history) associated with ecological and/or biogeochemical functions (resource acquisition, predator avoidance, and metabolite production) [1]. The purpose of this division is to simplify

community analyses and aid in the model building associated with climate change impacts on the biogeochemical and ecological components of oceans [2–4]. PFTs are usually defined using a combination of high phylogenetic grouping and functionality (for example, silicifying diatoms, mixotrophic dinoflagellates, nitrogen fixing and non-nitrogen fixing cyanobacteria, and calcifying coccolithophorids) or cell size [1,5–7]. Although size classification is not based on functional criteria, there is an alignment of the functional roles of phytoplankton with the size categories and the environmental niches (biogeochemical provinces) they occupy [1,8,9]. Altogether, the use of the PFT concept varies according to the scientific objectives of a given study and the observational capabilities available or required to tackle them [1].

Over the past two decades, the development of remote sensing techniques to identify PFTs has increased due to the wider spatial and temporal scales over which they can be used compared to in situ measurements. The most common satellite PFT approaches can be classified into abundance (size classes) and spectral (bio-optical) techniques [1,10,11]. Abundance-based approaches use satellite chlorophyll-a (Chl-a) as an input and, by using that, they exploit the largest signal in water-leaving radiance to extract the variability due to PFTs out of Chl-a. Spectral approaches are based on optical properties such as radiance, absorption, and/or backscattering spectra of phytoplankton, the variation of which is linked to community structure and pigment composition [10]. Different algorithms for the spectral-based discrimination of PTFs include the decomposition of absorption spectra, derivative analysis, inversion modeling, or Empirical Orthogonal Function (EOF) decomposition [11–16]. The validation of all these algorithms is a critical task that depends heavily on in situ measurements with sufficient spatial and temporal coverage [17]. Overall, the development and validation of PFT methods rely on HPLC (High-Performance Liquid Chromatography) pigment-based proxies of taxonomic composition or size structure and there is a clear need to complement these validations with supplemental datasets, including flow-cytometry, microscopy, and size-fractionated estimates [10,18,19].

Most algorithms perform similarly well over large gradients of co-varying bio-optical properties and reproduce expected trends in the global distribution of PFTs but not at smaller scales of variability [8]. Since different PTF algorithms use distinct approaches, datasets, and validation metrics, the evaluation of their performances requires comprehensive inter-comparisons using the same validation data and needs to consider errors associated with each one of them [10,20,21]. Thus, an appropriate choice of a PFT algorithm will depend on the scientific objectives and the observational capabilities available to validate its performance. PHYSAT is a spectral-based method that relies on the principle that spectral changes in anomalies of normalized water-leaving radiance (nLw) are due to changes in the dominant PFT [22]. This method was first developed for Case 1 waters using an empirical comparison of global SeaWiFS imagery with in situ data collected by the GEP&CO cruises (Geochemistry, Phytoplankton and Color of the Ocean [23]) along the North Atlantic, the Caribbean Sea, and the South Pacific. In its latest version, PHYSAT uses MODIS-Aqua data and discriminates between six phytoplankton groups (Nanoeukaryotes, *Prochlorococcus*, *Synechococcus*, Diatoms, *Phaeocystis*-like, and coccolithophorids) [24]. The current calibration of the PHYSAT algorithm has been updated using additional in situ measurements [25].

PHYSAT removes the dominant effect of Chl-a in the nLw spectrum to highlight the reflectance spectral anomaly generated by accessory pigments (that is, it exploits second-order anomalies of reflectance spectra) [22]. This proceeding is performed using an empirical reference model of nLw (nLw^{ref}) containing the average nLw at different Chl-a concentrations, based on a matchup between satellite records and in situ measurements. Since most of GEP&CO data come from open ocean sites with low Chl-a concentrations (<3 $mg \cdot m^{-3}$), the application of nLw^{ref} may not be appropriate in high Chl-a environments such as coastal and upwelling systems. In addition, coastal environments usually show a higher content of suspended inorganic matter and colored dissolved organic matter (CDOM), which could bias the PFT discrimination [10,25]. In the case of mid-latitude upwelling systems, most of the observed CDOM is transported from the subsurface by the same vertical Ekman fluxes that supply

nutrients and increase biological production in the surface layer; therefore, CDOM could increase and covary on similar spatial and temporal scales as phytoplankton biomass [26]. To minimize the effect of CDOM, PHYSAT can be adapted by the development of regional nLw^{ref} models. Recently, a regionalized version of PHYSAT was developed for the Mediterranean Sea (PHYSAT-Med) based on the optical retrievals of the phytoplankton assemblages and their succession patterns in this basin [27]. Thus, it is expected that the development of new regional bio-optical models will improve the capabilities of PHYSAT.

In the present study, we propose and test an alternative method to improve the identification of PFTs done using the PHYSAT approach. In theory, this new approach does not require the construction of regional reference models because it allows for an adequate PFT identification in both low and high Chl-a environments. The improved method, called PHYSTWO (PHYSAT based on second-order spectral modes), relies on the assumption that the dominant effect of Chl-a in the nLw spectrum can be effectively isolated from the signal of accessory pigments—biomarkers of different PFTs—by using EOF decomposition. To test and validate this approach, the PHYSTWO performance is compared against that of PHYSAT based on a regional nLw model for upwelling conditions using global datasets, as well as in-situ observations, from the coastal upwelling system (coastal and coastal transition zones) off central Chile.

2. Data and Methodology

2.1. Satellite Data and In Situ Measurements

Daily level-2 products of the MODIS-Aqua mission with a spatial resolution of 1 km were obtained from the OceanColor Web (http://oceancolor.gsfc.nasa.gov) for the region off central Chile (72–76°W and 35–38°S). The MODIS-Aqua products used were the Chl-a concentration estimated by the standard algorithm OC3, the aerosol optical thickness (AOT) at 869 nm, and the remote sensing reflectances (*Rrs*, in sr^{-1}) at 412, 443, 469, 488, 531, 547, and 555 nm. The *Rrs* were converted to normalized water-leaving radiances (nLw), adding the respective MODIS-Aqua spectrally convolved values of nominal solar irradiance (*Fo*, in $W{\cdot}m^{-2}{\cdot}\mu m^{-1}$) at each wavelength:

$$nLw = Rrs(\lambda) * Fo(\lambda) \quad (1)$$

The accuracy of PFT retrievals from both methods (PHYSAT and PHYSTWO) was assessed by comparing these with in situ data collected from different sources. Some of these data were obtained from in vivo spectral fluorescence profiles using a submersible spectro-fluorometer—the FluoroProbe (bbe Moldaenke GmbH, Kiel, Germany). This profiling instrument measures fluorescence emission at ~680 nm in response to excitation by light emitting diodes (LEDs) centered at ~370, 470, 525, 570, 590, and 610 nm. Its software provides estimates of CDOM or "yellow substances" and of the Chl-a concentration associated with the four PFTs derived from their fluorescence excitation spectra: green algae (Chlorophyta and Euglenophyta), brown algae (Bacillariophyta, Chrysophyta, and Dinophyta), blue algae (Cyanophyta), and red algae (Cryptophyta) [28]. In our study, the instrument did not register Cyanophyta because of a sensitivity problem (too high of a threshold level for its detection). Eighteen FluoroProbe profiles were obtained off Concepción (~36.5°S) as part of a cruise performed on an upwelling frontal area (PHYTOFRONT cruise) during 4–6 February 2014 (Figure 1). Six additional FluoroProbe profiles came from monthly samplings at a shelf time-series station (St. 18; 73°W and 36.3°S) during 2013. The CDOM data from the FluoroProbe profiles were used to assess its vertical distribution in the water column and its potential interference with the PFT retrievals.

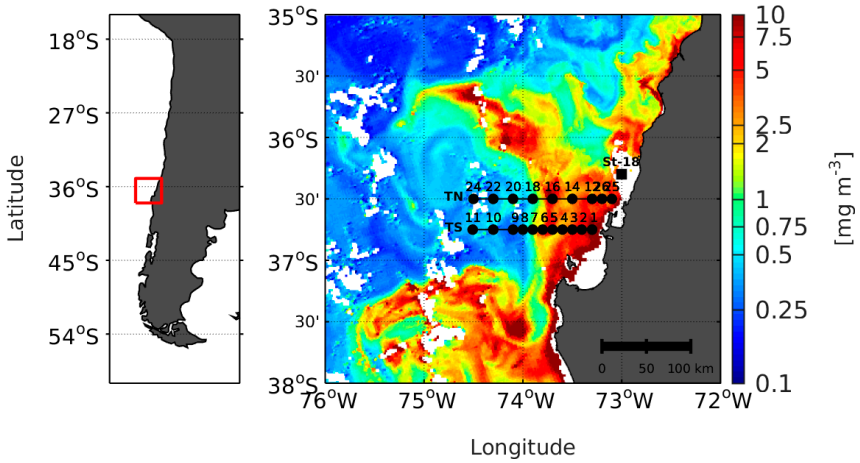

Figure 1. The mean daily chlorophyll-a (Chl-a) concentration derived from MODIS-Aqua for the dates of the PHYTOFRONT cruise (4–6 February 2014). The black lines and dots indicate the two transects and the stations sampled. The square symbol indicates the location of the time-series station St. 18.

Since the PFTs obtained with the FluoroProbe are not the same as those defined by PHYSAT, we used additional data collected during the PHYTOFRONT cruise to develop equivalences. These data were derived from the analyses of discrete water samples for PFT identification and enumeration using epifluorescence and inverted microscopy, as well as size-fractionated Chl-a and estimations of carbon biomass [29]. The FluoroProbe data were considered to represent two PHYSAT groups: the Nanoeukaryotes (including mostly flagellated green and red algae in pico-, nano-, and micro-size fractions) and the Diatoms (brown, non-flagellated algae, including nano- and micro-size fractions). PHYSAT originally defined Nanoeukaryotes as a group of small, flagellated algae containing characteristic carotenoids [24]. In the coastal region off Concepción, photosynthetic flagellates (including Chlorophyta and Cryptophyta) have been recognized as the largest contributors to the biomass of small phytoeukaryotes [29–32]. Hence, we conclude that the FluoroProbe Chl-a contributed by Chlorophyta and Cryptophyta (GA+C) represents the PHYSAT Nanoeukaryotes in this study. In addition, the FluoroProbe and PHYSAT Diatoms (Bacillariophyta) data can be considered as equivalent since this group is usually a dominant component of the microphytoplankton in the coastal zone during the spring–summer period [33]. FluoroProbe data from the upper 20 m layer for these two PHYSAT PFT (Nanoeukaryotes and Diatoms) were averaged and expressed as a percentage of the total Chl-a to compare their distribution with that derived from satellite observations (Figure 2).

For the remaining four PHYSAT PFTs, a direct comparison with in-situ data was not possible. During the PHYTOFRONT sampling, *Synechococcus* had low abundance/biomass and a lack of *Prochlorococcus* cells was presented [29]. Furthermore, no data on Coccolithophorids and *Phaeocystis* were obtained in this cruise. However, satellite estimates of *Phaeocystis* were compared using the available databases (BODC: British Oceanographic Data Centre; OBIS: Ocean Biogeographic Information System; OCB DMO: Ocean Carbon and Biogeochemistry Coordination and Data Management Office; Pangea: Data Publisher for Earth and Environmental Science; WOD09: World Ocean Database 2009; and US JGOFS: US Joint Global Ocean Flux Study) compiled by the MARine Ecosystem DATa (MAREDAT) initiative [34]. The dataset used here (141 samples) was restricted to samples collected during cloud-free days following the launch of the MODIS-Aqua mission (July 2002). Most of these correspond to coastal time-series stations. The *Phaeocystis* biomass data were transformed to Chl-a concentration values assuming a C:Chl-a ratio equal to 60:1 [35].

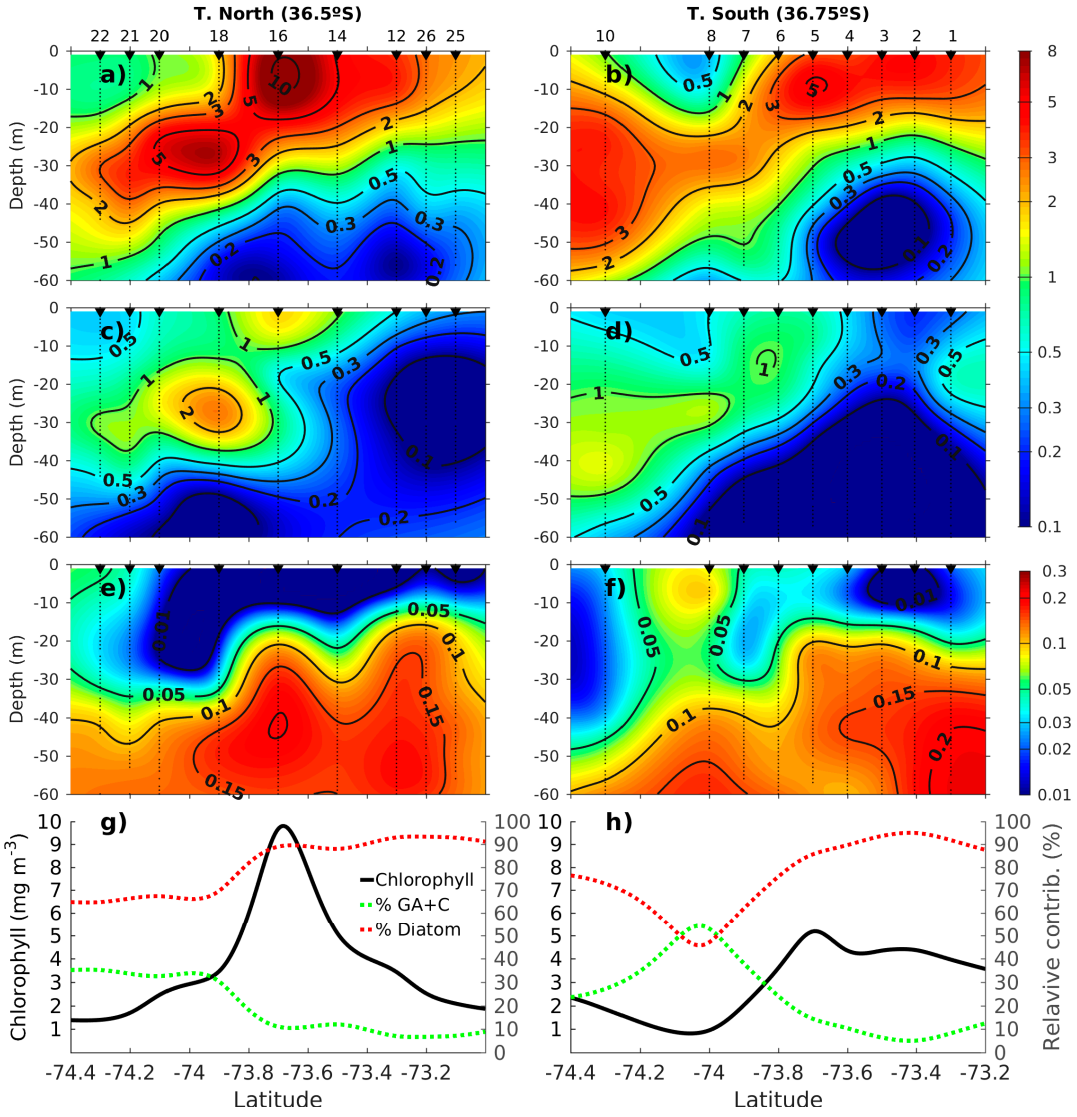

Figure 2. In situ FluoroProbe profiles during the PHYTOFRONT cruise: north and south transects (see Figure 1). The Chl-a concentration (color scheme; mg m^{-3}) associated to: Diatom (**a**,**b**); and Green Algae and Cryptophyta (GA+C; (**c**,**d**)); (**e**,**f**) color dissolved organic matter (CDOM) relative concentration; and (**g**,**h**) the average Chl-a concentration (black line) in surface waters (first 20 m depth) and the relative contribution of Diatom (red dashed lines) and GA+C (green dashed lines) to surface Chl-a.

2.2. The PHYSAT Regional Model for the Coastal Upwelling System Off of Central Chile

To compare the PHYSTWO and PHYSAT retrievals, an adapted version of PHYSAT was developed using a new regional reference model (nLw^{upw}) for the coastal upwelling region off central Chile. This model was used instead of the standard nLw^{ref} to remove the Chl-a spectral contribution from the nLw retrieved by the satellite since Chl-a has the dominant (first order) contribution of the retrieved signal from color sensors in open oceans [36,37]. The regional nLw^{upw} model is based on the mean of the daily MODIS-A nLw during the three months of the main upwelling activity (January–March 2014), using 1490 Chl-a concentrations in the range between 0.1 and 15 mg·m^{-3} (Figure 3), following the approach described in Reference [27].

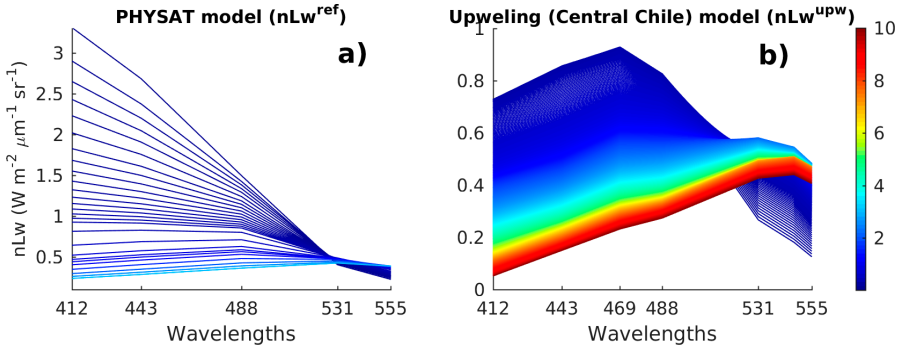

Figure 3. (**a**) An empirical reference model nLw^{ref} of PHYSAT [22] with the mean nLw radiances at 300 Chl-a concentrations (range: 0.1–3 mg·m^{-3}, every 0.1 mg·m^{-3}); and (**b**) a regional reference model nLw^{upw} with the mean nLw radiances at 1490 Chl-a concentrations (range: 0.1–15 mg·m^{-3}), based on the MODIS-A nLw data from the coastal upwelling region off central Chile (35–38°S and 72–76°W) during the upwelling season (January–March) of 2014. The Rrs retrievals where nLw(555) was >1.3 W·m^{-2}·μm^{-1}·sr^{-1} and the aerosol optical thickness (AOT) was >0.15 were excluded to reduce the presence of biased values arising from high concentrations of suspended sediments or errors in the atmospheric correction.

In the PHYSAT approach, the spectral removal of the Chl-a signal is achieved by dividing the observed nLw by the nLw^{ref} (or nLw^{upw}) to obtain the radiance anomalies (*Ra*) at each wavelength:

$$Ra(\lambda) = nLw(\lambda)/nLw(\lambda)^{upw}. \quad (2)$$

Alvain [22] showed that every PFT group can be associated with a specific *Ra* produced by the signal of their non-chlorophyll pigments. The obtained *Ra* are then compared with the acceptable ranges of *Ra* (Figure 4) for each PFT [24] to assign a dominant PFT at each pixel. Figure 5 displays the PFT fields obtained using nLw^{ref} and nLw^{upw}. The differences between the estimates of both models, as well as their accuracy, are discussed in Section 3.1.

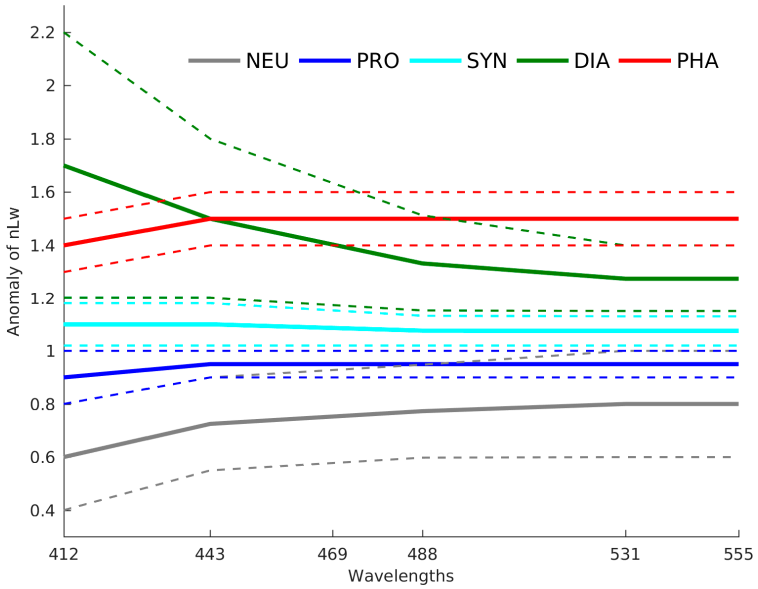

Figure 4. The acceptable PHYSAT ranges of the nLw anomalies (dashed lines) for each phytoplankton functional type (PFT) [24]. The continuous bold lines represent the mean values for each group. Abbreviations: COB: Coccolithophorids bloom; PHA: *Phaeocystis*-like; DIA: Diatoms; SLC: *Synechococcus*; PRO: *Prochlorococcus*; and NEU: Nanoeukaryotes.

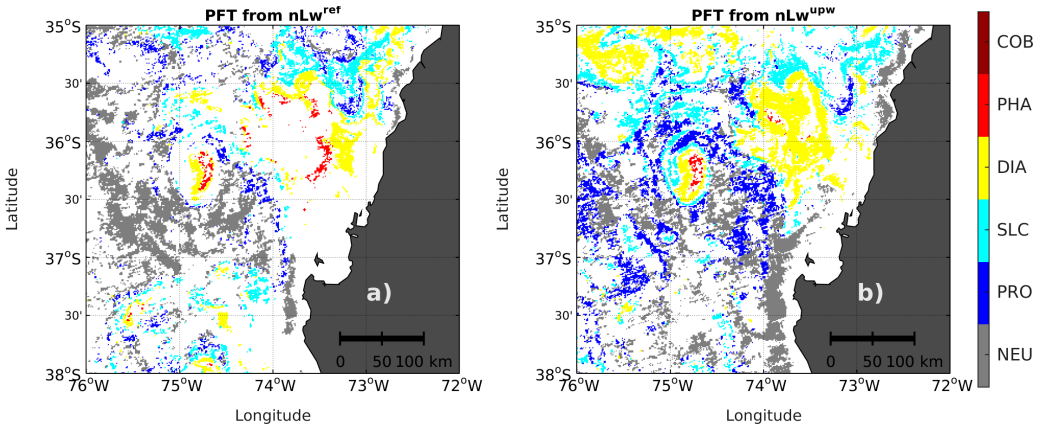

Figure 5. The PFT distribution calculated from a 3-day average of MODIS-A nLw data (4–6 February 2014), using: (**a**) the PHYSAT standard reference model (nLw^{ref}); and (**b**) the PHYSAT regional reference model (nLw^{upw}) for the coastal upwelling and transition zones in the region off central Chile. Abbreviations: COB: Coccolithophorids bloom; PHA: *Phaeocystis*-like; DIA: Diatoms; SLC: *Synechococcus*; PRO: *Prochlorococcus*; and NEU: Nanoeukaryotes.

2.3. Singular Value Decomposition (SVD) in Spectral Orthogonal Modes: The PHYSTWO Approach

The PHYSTWO method assumes that the Chl-a spectral signal, beyond being the dominant signal in the nLw retrieved by color satellites, is also independent (or not-correlated) from the signal of the other non-chlorophyll pigments (such as the ones used by PHYSAT to assign PFTs) or even from the signal of other substances such as CDOM in the surface ocean. From this perspective, it is expected that the Chl-a spectral signal can be separated from that of other pigments in the form of a dominant variability mode in an EOF analysis. An EOF analysis is a decomposition of a signal or dataset in terms of orthogonal modes which represent different (and, in theory, uncorrelated) fractions of data variability [38]. Usually, the first mode accounts for the largest fraction of variability. In nLw data from Case 1 (offshore) waters, this variability is expected to be associated to the Chl-a signal [36,37]. The remaining modes account for lower and uncorrelated fractions of variability and, therefore, they could contain the spectral signal of non-chlorophyll pigments that can be used to identify PFT assemblages.

PHYSTWO performs the EOF decomposition in a two-dimensional space–nLw matrix N (Equation (A2)). This matrix is a normalized version of the R matrix (Equation (A1)) which is built by reordering the satellite nLw fields, as is detailed in Appendix A. The N matrix is decomposed into three matrices using an SVD approach:

$$N = U * S * V^T. \tag{3}$$

The new matrix U has dimensions $P \times 7$ (P rows $\times$ 7 columns) and contains the spatial information of the N matrix. S is a diagonal matrix of dimensions 7×7 and V is a square matrix of dimensions 7×7 containing the spectral information. As in a typical SVD, the singular values on the diagonal S matrix are proportional to the explained variance contained in each mode m:

$$\%VarExp_m = \frac{S_m^2}{\sum_{i=1}^{m} S_i^2} * 100. \tag{4}$$

The SVD operates in a similar way on the R matrices constructed with the nLw data from different dates or geographic regions. In general, the first two resulting modes always account for ~99% of the explained variance ($m1$ ~60% and $m2$ ~39%) and the remaining 1% is scattered into the other five variability modes. The spatial pattern of the first two modes (U_1 and U_2) is recovered by rearranging

the dimensionless values of the first two columns of the U matrix in an original longitude–latitude matrix (Figure 6). A detailed inspection shows a close relationship between U_1 and the Chl-a field (Figure 1), whereas U_2 shows a good spatial agreement with the PFT retrievals obtained by PHYSAT (Figure 5). Therefore, U_2 may contain the spectral signal of non-chlorophyll pigments useful for PFT identification (see Sections 3.2 and 3.3). However, U_2 by itself does not provide information on the dominant PFT. Hence, it is not possible to establish reference values for each PFT (such as the PHYSAT acceptable ranges in Figure 4) in the U_2 field since dimensionless values obtained in the U, S, and V matrices can change depending on the numerical properties of the R matrix, the architecture of the computer hardware, and the numerical precision of the operating systems and programs used. For these reasons, SVD does not have a unique numerical solution. Any reference values must be incorporated before the SVD together with the values of the matrix R so that they undergo the same process of normalization and SVD that the R values have to ensure that they have comparable magnitudes in the resulting U-fields. The reference values must be organized within a synthetic matrix that can be attached to the R matrix without affecting the SVD.

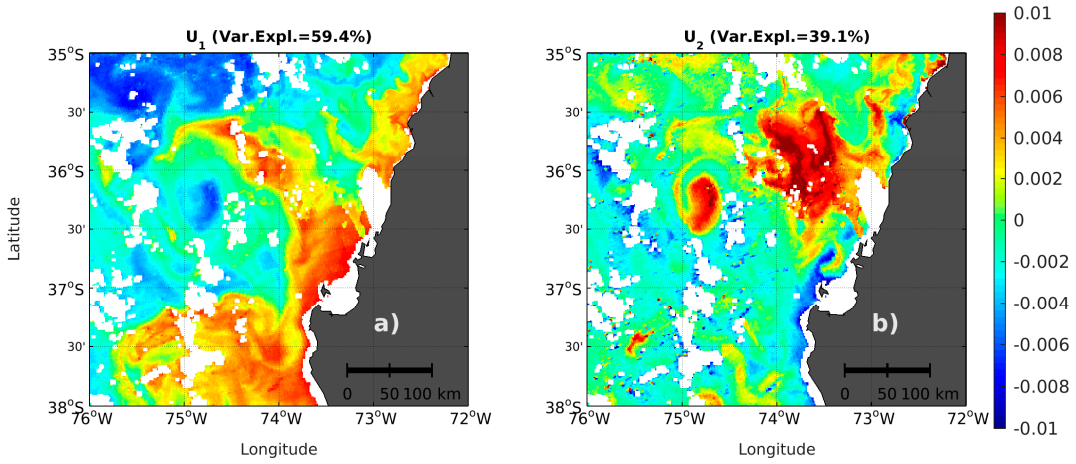

Figure 6. The spatial pattern of the: first U_1 (**a**); and second U_2 (**b**) orthogonal modes derived from the SVD of a space–nLw matrix, composed by the average MODIS-A nLw data from the 4–6 February 2014.

2.4. The Synthetic Matrix of Typical nLw Values for PFT Categories and PFT Estimation from Orthogonal Models

Obtaining a reference nLw matrix useful for PFT identification is not an easy task since it requires observational information about abundance, biomass, and the spectral signal generated by different concentrations of several phytoplankton groups in waters with different optical properties. This type of information is generally scarce because it requires intense field measurements and laboratory analyses [34]. A first version of a synthetic matrix *Rs*, containing reference values for PFT, can be obtained from the PHYSAT acceptable ranges of radiance anomalies (Figure 4) [22,24,27] following the procedure detailed in Appendix B. This *Rs* matrix contains the expected radiances of each dominant PFT under different Chl-a concentrations. Therefore, it is a concentration–nLw matrix whose values have the same units ($W \cdot m^{-2} \cdot \mu m^{-1} \cdot sr^{-1}$) of the R matrix built with satellite nLw data. Although R is a space–nLw matrix and does not have the same size of Rs, both matrices have the same number of columns as they share the same wavelength frequencies. Therefore, it is mathematically possible to concatenate the *Rs* matrix after the R matrix to obtain a new combined matrix *RRs*:

$$RRs = \begin{pmatrix} R \\ R_S \end{pmatrix}. \quad (5)$$

The concatenation is a common variation of the EOF method used in the analysis of the co-variability (or joint variability) of two or more fields at a time [38]. Here, the concatenation

was performed as a way to provide reference values before the SVD. As was done for the *R* matrix, the *RRs* matrix is also standardized by removing the mean and standard deviation (Equation (A2)) to obtain the *NNs* matrix. The SVD of the standardized matrix is achieved through:

$$NNs = U * S * V^T. \quad (6)$$

This decomposition results in a *U* matrix, containing in its last 1800 rows, the dimensionless reference values of PFT at different Chl-a concentrations. As for the decomposition of the *N* matrix, the first two modes of the *NNs* decomposition accounts for most of the explained variance (M1 = 58.7% and M2 = 39.8%).

In the dimensionless plane formed by the first two spatial modes U_1 and U_2, we can locate the orthogonal values resulting from the matrix *R* decomposition and the typical values for the PFT derived from the decomposition of the synthetic matrix *Rs* (Figure 7). The typical PFT values appear distributed from the lower right to the upper left corner in Figure 7. In this same direction, the Chl-a concentration increases and therefore, the typical values at low (high) concentrations are located in the lower-right (upper-left) sector. It is noticeable that the typical values of all PFTs are distributed almost parallel to each other, without any overlap between them. This is an inherited characteristic of the acceptable ranges of the Ra matrix which helps to avoid a double assignment of PFTs.

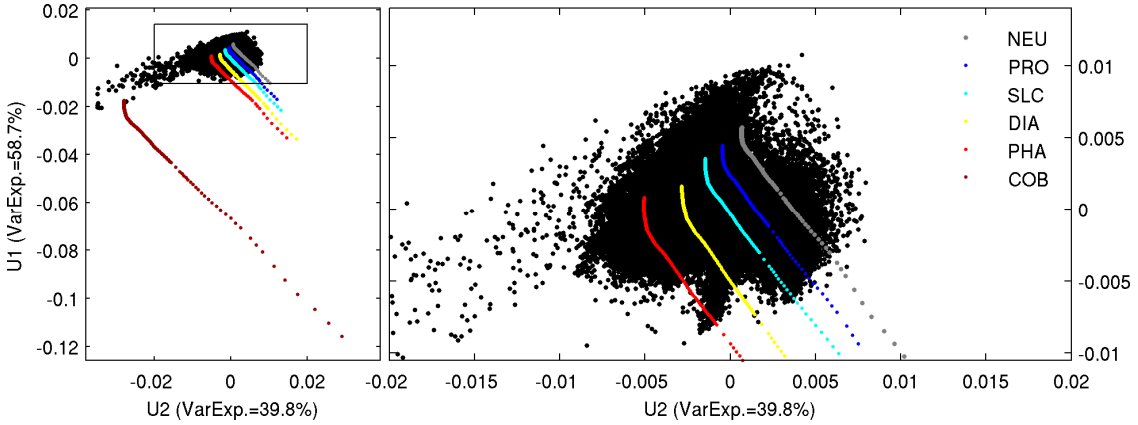

Figure 7. The dimensionless U_1–U_2 plane derived from the singular value decomposition (SVD) of *NNs* matrix. The colored symbols represent the location of the typical values (synthetic matrix *Rs*) of the PFTs at different Chl-a concentrations; the black dots correspond to values derived from the observations at each pixel in the *R* matrix. Abbreviations are the same as in Figure 5.

The lines formed by typical values in the U_1–U_2 plane can be considered as PFT orthomodels. The assignment of the PFTs corresponding to each point in the R matrix is easily achieved using the Euclidean distance to the orthomodels values. The PFTs assigned to any point will be that of the orthomodel with the closest typical value in the U_1–U_2 plane, as shown in Figure 8a. By rearranging the assignments in a longitude–latitude space, the first and unadjusted PHYSTWO estimation field is obtained (Figure 8b). A schematic diagram summarizing the PHYSTWO method is shown in Figure 9.

The unadjusted PFT estimation of PHYSTWO does not fully agree with the in-situ observations made during the PHYTOFRONT cruise (see Section 3.4). Some erroneous assignments are due to the fact that current orthomodels have no values in the region of high Chl-a concentration in the U_1–U_2 plane (since the typical values in the Rs matrix come from waters with relatively low Chl-a concentrations <3 mg·m^{-3}; GEP&CO cruises). Additionally, some PFT reference values are out of the range of their known environmental preferences of background Chl-a concentration. It is possible to improve the PHYSTWO estimations by directly fitting the orthomodels in the U_1–U_2 plane by using additional information (field measurements or available databases) as described in the next subsection.

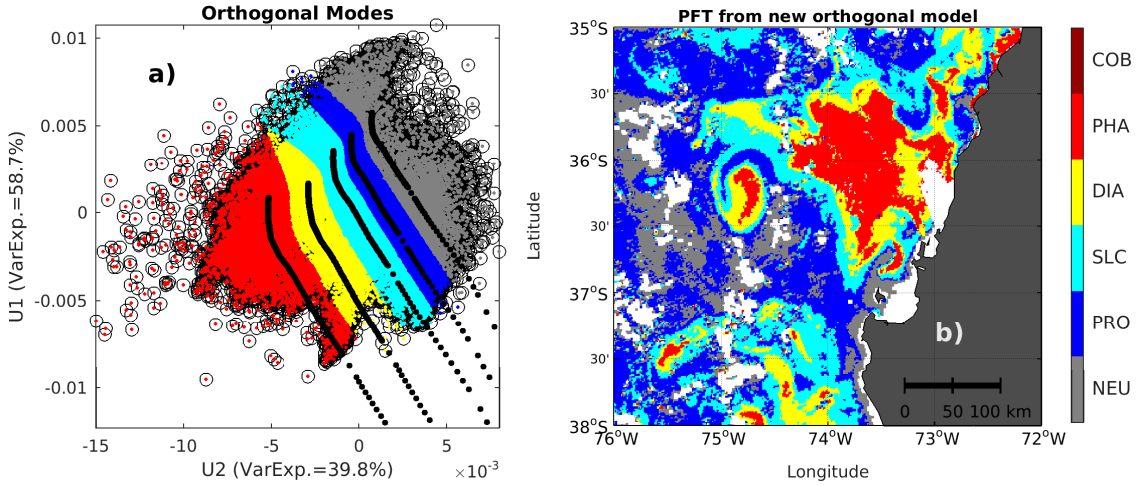

Figure 8. (**a**) The dimensionless U_1–U_2 plane derived from the joint SVD of the synthetic matrix *Rs* and an *R* matrix composed by the average MODIS-A nLw data during 4–6 February 2014. The colors represent the respective PFTs assigned to each dot (pixel) considering their closeness to the unadjusted PFT orthomodels (black dots) obtained from the acceptable ranges proposed by PHYSAT. (**b**) The unadjusted PFT estimation of PHYSTWO by rearranging the assignments of (**a**) in a longitude–latitude plane. Abbreviations are the same as in Figure 5.

2.5. The Adjustment of PFT Orthomodels with In Situ Data

2.5.1. Diatoms and Nanoeukaryotes

Figure 10a shows the U_1–U_2 plane obtained from the SVD of an Rs matrix together with an R matrix composed by the matchups with the satellite nLw values corresponding to the place and date of the samples during the PHYTOFRONT cruise (16 stations) and St. 18 (four stations). Each matchup was obtained as the mean nLw value in a 3 × 3 pixels box (~9 km^2) around the latitude and longitude of any station for the same date. The size of the circles represents the relative abundance of Diatoms and Nanoeukaryotes based on FluoroProbe profiles. In this plane, Diatoms are most abundant at stations located in the upper left corner, associated with higher Chl-a concentrations (3.5–10.5 mg·m^{-3}). These stations are located beyond the Diatom orthomodel that considers Chl-a concentrations <3 mg·m^{-3}. To fit the diatom orthomodel, an extension towards the region of higher Chl-a concentrations is required. This is achieved by incorporating new typical values following the path determined by Diatom proportion values of up to 95% in the interpolated relative abundance field, as represented by the orange dots in Figure 10b.

Figure 10 also indicates that the Nanoeukaryotes were relatively more important in stations located towards the right central and upper sectors in the U_1–U_2 plane. However, the upper end of the Nanoeukaryotes orthomodel extends to the region where our observations suggest a greater dominance of Diatoms. Therefore, the fit for Nanoeukaryotes was done by excluding the top points of the orthomodel where the Nanoeukaryotes proportion was <45% (white dots, Figure 10c). Unfortunately, the PHYTOFRONT data were not sufficient to modify the orthomodels in regions of moderate to low Chl-a concentrations. These adjustments should be addressed in future studies covering information from the mesotrophic and oligotrophic regimes. However, considering that most of GeP&CO observations come from open ocean regions, we expect that the current values of the orthomodels in this range of Chl-a concentrations do not require significant changes.

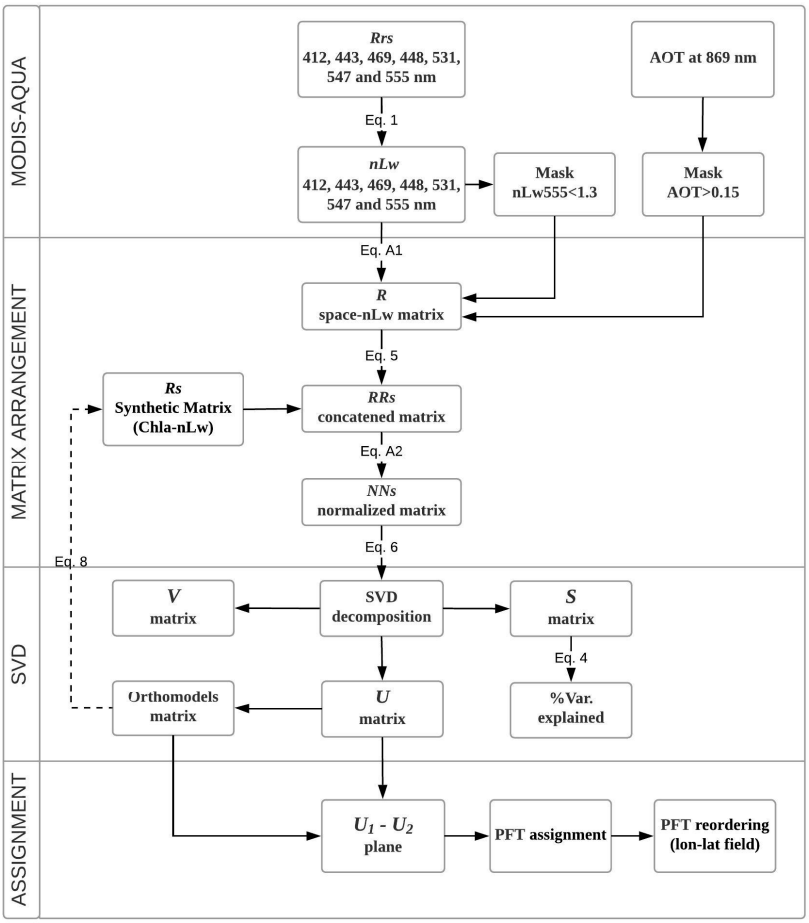

Figure 9. The schematic diagram summarizing the PHYSTWO method. The dashed arrow represents the process of orthomodels adjustment and their incorporation into the synthetic matrix.

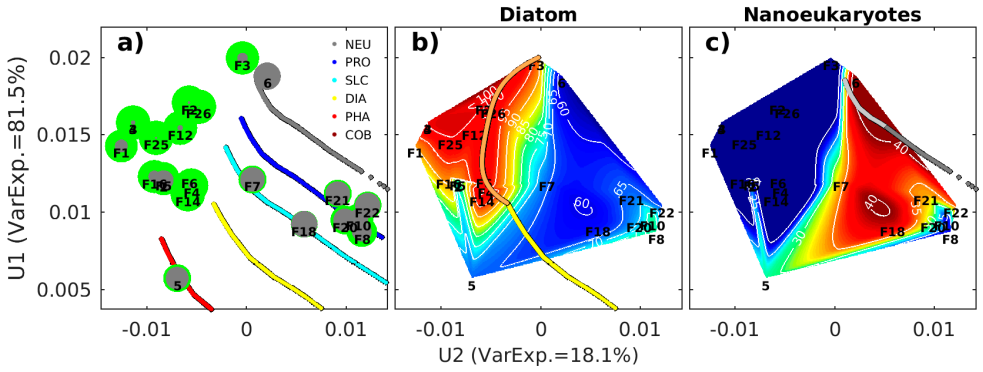

Figure 10. The dimensionless U_1–U_2 plane obtained from the joint SVD of the synthetic matrix *Rs* and an *R* matrix composed by the matchups of nLw data corresponding to in situ measurements performed during the PHYTOFRONT cruise and in St. 18. (**a**) The relative concentration (percent of total Chl-a concentration) of Diatoms (green circles) and Nanoeukaryotes (grey circles) derived from FluoroProbe measurements associated with Diatom and Green Algae plus Cryptophyta, respectively. In (**b**,**c**), the colored area represents the relative concentration interpolated for Diatoms and Nanoeukaryotes, respectively, with blue (red) colors corresponding to low (high) values. The orange dots in (**b**) denote the new addition to the orthomodel for Diatoms (yellow line); the light gray dots in (**c**) denote the points excluded in the orthomodel for Nanoeukaryotes (grey line). Abbreviations are the same as in Figure 5; F-coding refers to the PHYTOFRONT stations in Figure 1.

2.5.2. *Phaeocystis*

The worldwide distribution of *Phaeocystis* is poorly understood but the locations that it is abundant in have been reported in the Southern Ocean southward of 70°S and in the North Atlantic northward of 50°N [34]. In the coastal upwelling system off Peru, the presence of *Phaeocystis globosa* has been reported during iron-fertilization experiments [39]. However, further in situ data on *Phaeocystis* in the coastal upwelling area off Chile are not available [40]. Despite this lack of information, retrievals from PHYSAT and PHYSTWO (Figures 5 and 9) suggest that this PFT may be present in our study region.

The SVD of the *R* matrix corresponding to the MAREDAT stations shows a core of high *Phaeocystis* abundance that coincides with the middle region (moderate Chl-a concentrations) of the *Phaeocystis* orthomodel. Since PHYSAT lacks direct evidence to constrain the acceptable ranges for *Phaeocystis*, these estimates were initially categorized as *Phaeocystis*-like [24]. The spatial coincidence between the high concentrations of *Phaeocystis* observed from MAREDAT data and the orthomodel suggests that the PHYSAT method made an adequate assignment of this PFT, at least in regions with intermediate background concentrations of Chl-a. Under this premise, the *Phaeocystis* orthomodel was fitted to just exclude values corresponding to pixels with high and low Chl-a concentrations in the area below 30% of the *Phaeocystis* abundance (dark red dots, Figure 11).

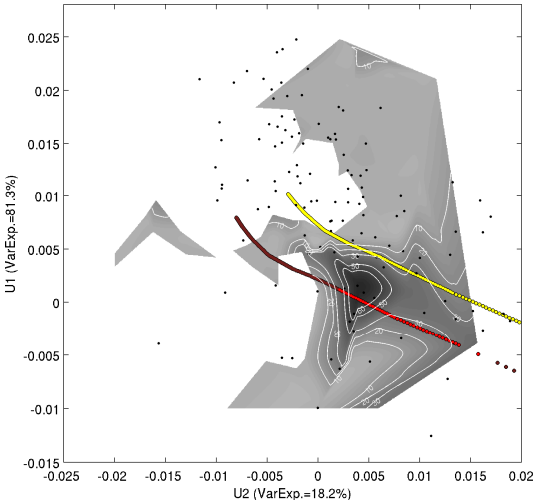

Figure 11. The relative concentration of *Phaeocystis* (gray tones) in the orthogonal U_1–U_2 plane, calculated from the SVD of an *R* matrix composed of 141 registers from the MAREDAT database in the period between 2002 and 2009. The orthomodels for *Phaeocystis* and Diatoms are shown as red and yellow dots, respectively. The bright red dots show the proposed adjustment for the *Phaeocystis* orthomodel.

2.5.3. *Prochlorococcus* and *Synechococcus*

Photosynthetic cyanobacteria have a wide distribution and are dominant in the ocean regions with very low Chl-a concentrations where low nutrient availability does not favor the dominance of other large-size PFTs. For this reason, the orthomodel extension towards the upper left corner of high background Chl-a concentration is not consistent with what we know about the ecological niches of the dominant cyanobacteria genera, *Prochlorococcus* and *Synechococcus* [41]. Using molecular techniques, Bibby et al. [42] have shown an alternation in the dominance of *Prochlorococcus* and *Synechococcus* in open waters related to the background concentration of Chl-a and nutrients. *Prochlorococcus* only appears dominant at Chl-a concentrations <0.35 mg·m^{-3}, while *Synechococcus* dominates at values >0.35 and <0.5 mg·m^{-3}. However, in regions with iron limitation, *Synechococcus* can extend its dominance to the 0.26 and 0.5 mg·m^{-3} range thanks to the greater efficiency of its photosynthetic system under stressed conditions compared to that of *Prochlorococcus*. Based on these results,

the *Prochlorococcus* orthomodel was restricted to the typical values associated with Chl-a concentrations between 0.1 and 0.3 mg·m^{-3} and the *Synechococcus* orthomodel to values between 0.26 and 0.5 mg·m^{-3} (Figure 12).

2.6. PFT Estimation with Adjusted Orthomodels

A new matrix of typical values (*Rt*) can be obtained with a reconstruction of the adjusted orthomodels. The fitted values in U_1 and U_2 were included in a new matrix $Ut_{1,2}$ and multiplied by their corresponding values in the *S* and *V* matrices to obtain the new synthetic *Rt* adjusted matrix:

$$Rt = Ut_{1,2} * S_{1,2} * V_{1,2}^{T} \quad (7)$$

After the mean and standard deviation (which were previously removed before the SVD) are added back, *Rt* becomes a matrix of nLw radiances expected for each PFT at different background Chl-a concentrations. Hereafter, this new matrix can be used to concatenate with other R matrices and, in this way, will allow for an adjusted PFT estimation. In our study case, PFT estimations using *Rt* (Figure 13) show a large agreement with the in situ measurements rather than the PHYSAT estimations (see Section 3.4). This new PHYSTWO method and the *Rt* matrix of typical adjusted values were implemented as a Matlab function and are freely available at https://www.dropbox.com/sh/1j7sn3uo7fa60r5/AABwub8DEg1YOqsICY_PAnkEa?dl=0.

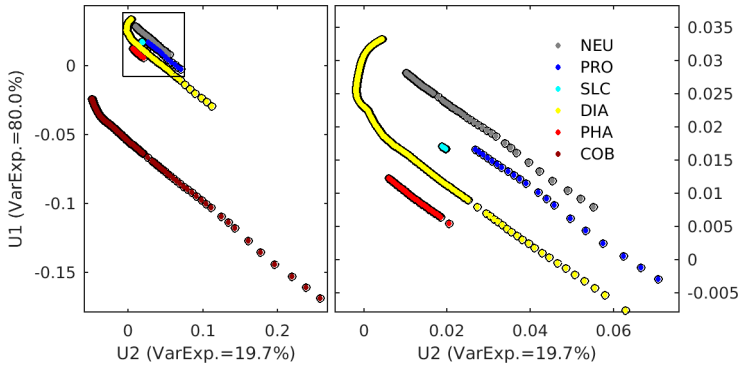

Figure 12. The fitted orthogonal models (first U_1 and second U_2 modes) for each PFT. The section on the right is an enlargement of the square area in the section on the left. Abbreviations are the same as in Figure 5.

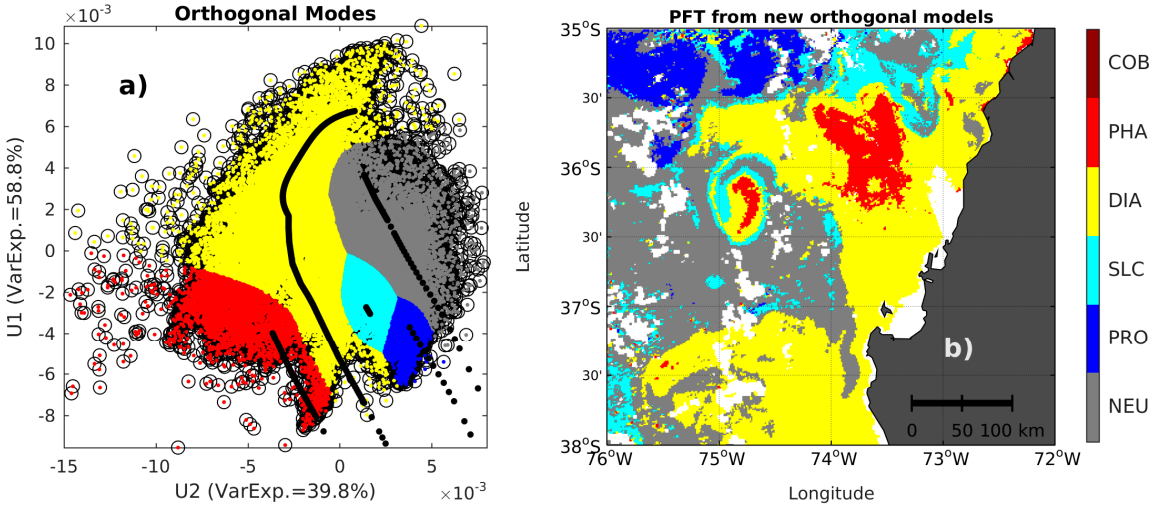

Figure 13. *Cont.*

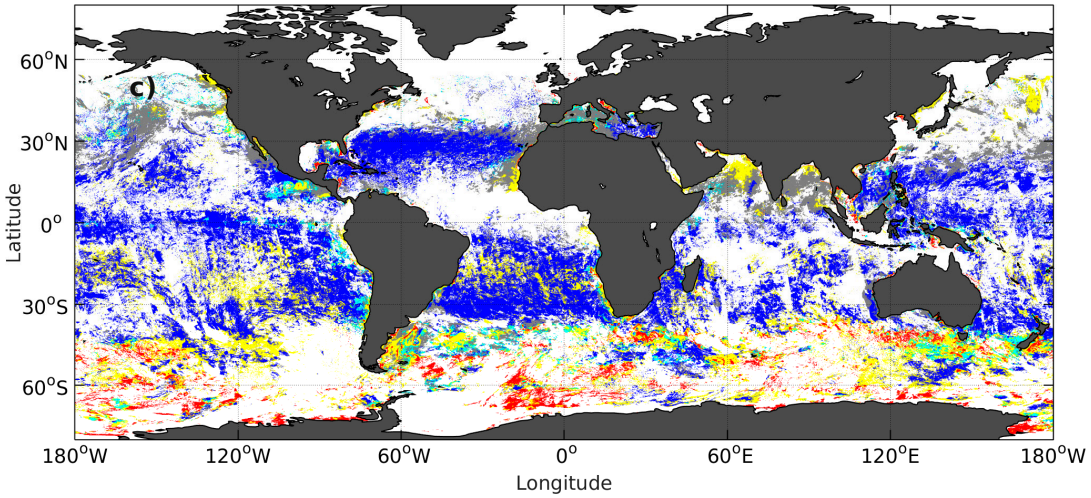

Figure 13. (**a**) The dimensionless U_1–U_2 plane derived from the joint SVD of the typical values matrix *Rt* and an *R* matrix composed of the average MODIS-A nLw data during 4–6 February 2014. The colors represent the respective PFT assigned to each dot (pixel) considering their closeness to the fitted PFT orthomodels (black dots) shown in Figure 12. (**b**) The adjusted PFT estimation of PHYSTWO were obtained by rearranging the assignments of (**a**) in a longitude–latitude plane. (**c**) A global view of the PFT estimation performed by PHYSTWO for the same dates. Abbreviations are the same as in Figure 5.

3. Results

3.1. Adaptation of PHYSAT to Upwelling Conditions

To obtain more accurate PFT estimates that could be used to compare and measure the performance of PHYSTWO, the performance of PHYSAT was improved by the development of a regional model nLw^{upw} for the coastal upwelling region off central Chile (35–38°S and 72-76°W). The new nLw^{upw} greatly expands the original PHYSAT Chl-a upper limit from 3 to 15 mg·m^{-3} and it allows PFT estimations at higher Chl-a concentrations than it is possible using nLw^{ref}. This is consistent with observations in the region where Chl-a concentrations >10 mg·m^{-3} are common in coastal waters during the upwelling season [33]. The resulting nLw^{upw} values were lower (about half) than the standard nLw^{ref} and displayed a dome pattern due to a central maximum at 469 nm (Figure 3b). This pattern is not unexpected since the shape and amplitude of the spectra are not unique to the global ocean but are highly dependent on the bio-optical environment of each region [25]. Furthermore, it is not possible to compare directly both reference models at a wavelength of 469 nm since this wavelength does not form part of the nLw^{ref} model originally proposed by Alvain et al. [22].

The distributions of the dominant PFTs based on the standard PHYSAT reference model (nLw^{ref}) and the regional model for the coastal upwelling and transition conditions (nLw^{upw}) during the dates of the PHYTOFRONT cruise are presented in Figure 5. The nLw^{ref} model (Figure 5a) produces PFT retrievals for only 26.9% of the valid pixels (pixels with valid nLw data and free of cloud interference), while over 70% of pixels remain without an assigned PFT, especially in the coastal zone where Chl-a exceeds 3 mg·m^{-3} (Table 1). The retrievals in the coastal upwelling zone suggest a dominance of Nanoeukaryotes. Diatoms, on the other hand, dominated a high Chl-a filament and an anticyclonic subsurface mesoscale eddy, structures which were observed in the transition from the coastal to the oceanic zone. *Phaeocystis* was observed along the filament's central axis. In the rest of the coastal and offshore areas, *Synechococcus* and *Prochlorococcus* were the dominant groups.

The PFT distribution using the nLw^{upw} model (Figure 5b) is similar to the one derived from nLw^{ref}. However, it provides a slight improvement regarding PFT retrievals (39.5%) in high Chl-a areas. One of the differences between the models was that the nLw^{upw} model showed a higher spatial coverage of Diatoms (10.3%) than the nLw^{ref} model (3.0%). Both PFT models (nLw^{ref} and

nLw^{upw}) yielded a greater dominance of Nanoeukaryotes in the coastal zone, an expected result during the relaxation of the upwelling conditions [30]. During the PHYTOFRONT cruise, a relaxation of upwelling was detected, and this might have led to a slightly higher dominance of the nanoplankton Chl-a fraction compared with the microplankton in the upper layer [29]. The coastal retrievals of pico-prokaryotic phytoplankton by PHYSAT are in clear disagreement with the environmental preference of *Prochlorococcus* for Chl-a < 0.35 mg·m^{-3} [42]. Besides, during the PHYTOFRONT cruise, low abundances of *Synechococcus* and no cells of *Prochlorococcus* were detected by flow-cytometry [29]. These results suggest that PHYSAT based on either the standard or the regional reference models may not be fully appropriate for the estimation of PFT distributions in waters with high Chl-a concentrations as those sampled during the PHYTOFRONT cruise.

Table 1. The percentage of phytoplankton functional types (PFT) retrievals obtained from the PHYSAT and PHYSTWO methods off central Chile for 4–6 February 2014.

Method	Retrievals	% Ret. [1]	NEU (%)	PRO (%)	SLC (%)	DIA (%)	PHA (%)	COB (%)
PHYSAT nLw^{ref}	18,944	26.9	16.1	3.1	3.3	3.0	1.4	0.0
PHYSAT nLw^{upw}	27,874	39.5	11.9	10.1	7.1	10.3	0.2	0.0
PHYSTWO unadj.	70,009	99.3	18.6	35.4	22.6	10.1	11.5	0.0
PHYSTWO adj.	70,009	99.3	37.7	8.9	8.6	38.4	5.7	0.0

[1] The percentage of retrievals in relation to the total number of pixels with valid nLw data and where satellite Chl-a estimation was possible.

3.2. First Spectral Mode and Satellite Chl-a

The PHYSTWO approximation relies on the assumption that Chl-a is the dominant and independent signal in the nLw retrieved by ocean color satellites and, therefore, it is susceptible to be separated as the first variability mode through an SVD. The SVD of the space–nLw matrix composed by MODIS-A nLw data off central Chile, on the dates of the PHYTOFRONT cruise (4–6 February 2014), resulted in seven variability modes. The first two modes together accounted for most of the explained variance (>99%). The spatial pattern of the first mode (U_1, VarExp = 79.9%, Figure 6a) displays a close similarity with that of the Chl-a field (Figure 1), with higher U_1 values in the coastal zone and in the filament north of the anticyclonic eddy. Indeed, Chl-a displays a high correlation with U_1 (r = 0.71) and does not show a significant correlation with U_2 (r = 0.05; Figure 7a,b). This high correlation suggests that the first mode contained the major part of the Chl-a spectral signal and, therefore, the back reconstruction of the first-order matrix M_1 must contain nLw values associated to chlorophyll reflectance. This reconstruction is easily achieved using Equation (8), where U_1, $V_{1,}$ and S_1 are the first column of these matrices:

$$M_1 = U_1 * S_1 * V_1^T. \quad (8)$$

Since the M_1 matrix contains the Chl-a signal of the phytoplankton community, this could be used to obtain a first-order reference model nLw^{M1} (Figure 14c) following the same proceeding as that used to obtain the regional upwelling model nLw^{upw} in Section 2.2. When these models are compared, it is possible to observe that nLw^{M1} has a close similarity to nLw^{upw}, in both shape and magnitude, showing the same spectral threshold at the 469 frequency, which determines a high correlation between both models (r = 0.91; Figure 14d). This spectral similarity corroborates that the M_1 mode—where the nLw^{M1} comes from—contains the spectral signal of Chl-a.

The small differences between nLw^{M1} and nLw^{upw} may be linked to other substances with a strong effect on the spectral nLw radiances such as suspended sediments or CDOM. The FluoroProbe profiles show an increase of CDOM with depth (Figure 2), leading to a low but significant inverse correlation with the vertical distribution of Chl-a (Table 2). This pattern has been previously reported as a typical condition in upwelling systems [26]. However, the low CDOM concentrations in the upper 20 m do not have significant relationships either with the MODIS-A Chl-a or with the first two

spatial modes. This suggests that the CDOM signal in the study region may represent a negligible contamination on the MODIS-A retrievals and in the subsequent PFT estimation based on the SVD of the R matrix, as performed by PHYSTWO.

Table 2. The correlation coefficients of the relationship between Colored Dissolved Organic Matter (CDOM), measured from FluoroProbe profiles (CDOM-FP, 0–60 m depth) and integrated in the upper 20 m layer (CDOM-20m), and the in-situ measured FluoroProbe Chl-a (Chl-FP), the satellite MODIS-A chlorophyll-a (Chl-sat), and the first two spatial modes obtained from the singular value decomposition (SVD) of a *R* matrix composed by MODIS-A nLw data in the region off central Chile (35–38°S, 72–76°W).

	Chl-FP	Chl-sat	U_1	U_2
CDOM-FP	−0.29 ^{S}	-	-	-
CDOM-20m	-	−0.14 ^{NS}	−0.11 ^{NS}	0.22 ^{NS}

^{S} The correlation is significant at the 0.05 level; ^{NS} Not significant.

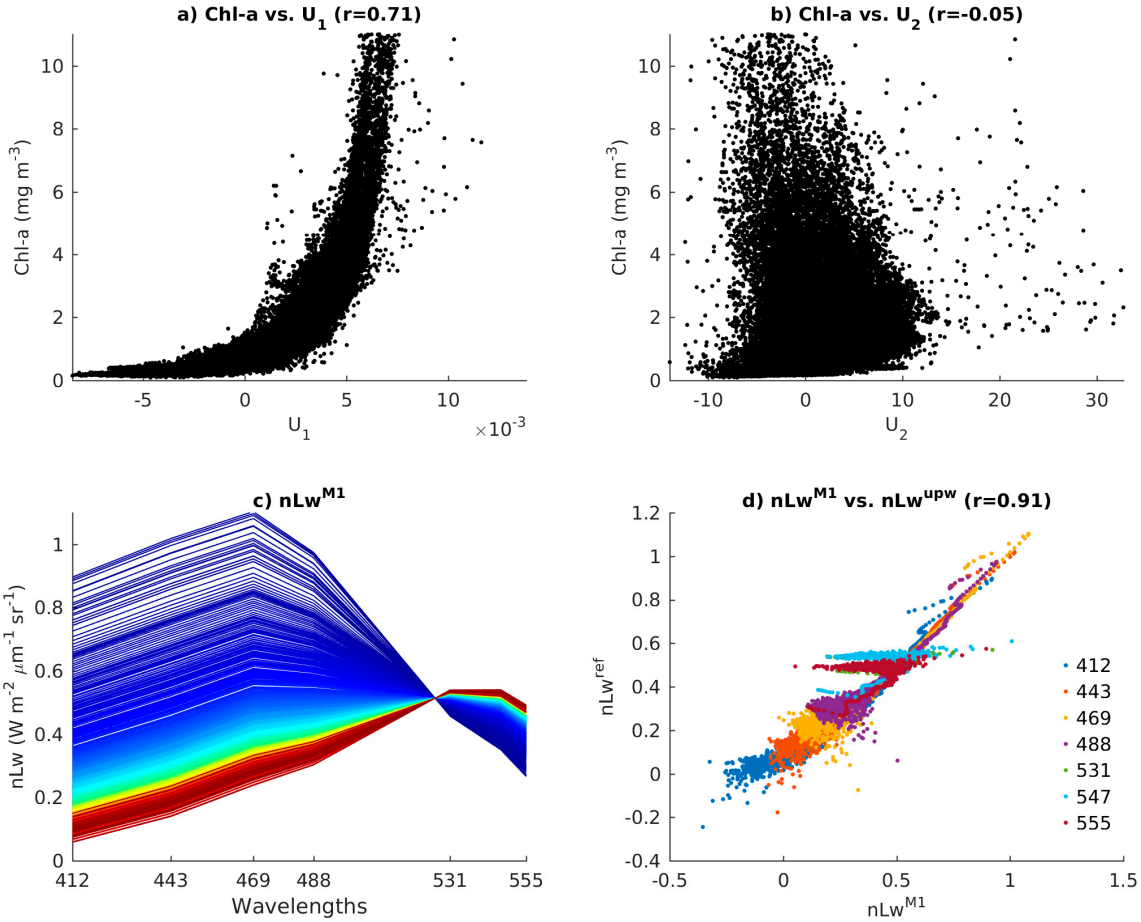

Figure 14. The relationship between MODIS-A Chl-a and the spatial pattern of the: first (**a**); and second (**b**) modes resulting from of the SVD of a *R* matrix composed by MODIS-A nLw data off central Chile (35–38°S, 72–76°W) on the dates of the PHYTOFRONT cruise (4–6 February 2014). (**c**) The regional reference model (nLw^{M1}) derived from the reconstruction of the first SVD orthogonal mode M_1. (**d**) The relationship between nLw^{M1} and nLw^{upw} for each wavelength.

3.3. Use of the Second Spectral Mode for Analyzing PFT Spatial Succession

The variability associated with the second spatial mode (U_2, VarExp = 39.1%) is, by definition, orthogonal and independent from U_1 variability. Since U_1 contains the variability associated with Chl-a

(see the previous section), we expect U_2 to contain the signal associated with non-chlorophyll accessory pigments that can be used as indicators of the PFT. Evidence of this can be found in the spatial pattern of U_2 (Figure 6b), which closely resembles the PFT field derived from PHYSAT (Figure 5). Comparing these two fields, we observe that high U_2 values in the filament zone correspond to regions dominated by Diatoms, while low U_2 values fall in regions dominated by Nanoeukaryotes and intermediate values by smaller phytoplankton. This spatial coherence is expected and relies on the fact that PHYSAT is also designed to highlight the effect of secondary pigments that are overshadowed by the Chl-a signal [22]. Unfortunately, there is no way to quantify the degree of agreement between U_2 and PFTs because U_2 is a continuous field (that is, not categorical like the PFT estimations of PHYSAT) and because our results show that PHYSAT estimations of PFTs that could be used in this comparison may not be appropriated for the region. Indirect evidence of this agreement is obtained from the fact that an accurate PFT identification based in the U_2 field is possible, as performed by PHYSTWO.

3.4. PFT Accuracy for the PHYSAT and PHYSTWO Methods

PFT retrievals obtained by the PHYSAT (using nLw^{ref} and nLw^{uwp} reference models) and PHYSTWO (using unadjusted and adjusted orthomodels) were compared with the relative abundance of Nanoeukaryotes and Diatoms observed along the PHYTOFRONT transects off central Chile (Table 3). The *Synechococcus* and *Prochlorococcus* groups were not dominant or present in these transects but they are expected to be present offshore, out of the direct influence of upwelling waters and where low nutrient conditions coincide with their known environmental preferences. Therefore, PFT retrievals for these groups were compared with the potential dominance they may have in areas defined by their ecological preferences in terms of Chl-a concentration (as a proxy of nutrient concentration) as reported by Bibby et al. [42]. For the remaining groups (*Phaeocystis* and Coccolithophorids), we were not able to perform any validation since neither in situ nor published data were available for the study region. However, considering that retrievals for *Phaeocystis* are obtained by both PHYSAT and PHYSTWO estimations, besides the fact that previous studies have reported its presence in upwelling system, it is expected that this group could be present in the region.

Table 3. The comparison between PFT retrievals with in situ and published data. Ret = the total number of retrievals on PHYTOFRONT transects. % Agr. = The percentage of retrievals on PHYTOFRONT transects that are in agreement with the in situ dominance of Diatoms and Nanoeukaryotes measured in the FluoroProbe profiles. % PAgr. = the percentage of retrievals that are in agreement with the areas of likely dominance by *Prochlorococcus* (where Chl-a concentration is <0.35 $mg \cdot m^{-3}$) and *Synechococcus* (where Chl-a is between 0.26 and 0.5 $mg \cdot m^{-3}$).

Method	NEU		DIA		PRO		SLC	
	Ret.(1)	% Agr.	Ret.	% Agr	Ret.	%PAgr.	Ret.	%PAgr.
PHYSAT nLw^{ref}	64	6.1	0	0.0	2169	10.3	2305	9.7
PHYSAT nLw^{uwp}	56	0.0	9	1.9	7111	13.8	5028	8.0
PHYSTWO unadj.	73	0.0	93	19.7	24,975	55.7	15,955	11.9
PHYSTWO adj.	174	81.6	327	68.9	6289	29.9	6051	22.1

Using the nLw^{ref} reference model, PHYSAT does not produce any Diatom retrievals for the PHYTOFRONT transects, although this group was the dominant component in 90.6% of the FluoroProbe profiles (Figure 2). There were 64 Nanoeukaryote retrievals for the transects but only 6.1% of these were observed as dominant in the FluoroProbe profiles. When the regional model nLw^{upw} is used, PHYSAT slightly increases the Diatoms retrievals and the percentage of agreement with FluoroProbe profiles to 1.9% but it does not produce any coincidences for Nanoeukaryote retrievals. The two PHYSAT models produced an important number of retrievals for *Prochlorococcus* and *Synechococcus*. However, only a low percentage (13.8% and 8.0%, respectively) of such retrievals coincided with their environmental preference, suggesting a misassignment for these groups.

PHYSTWO estimations produce retrievals over 99.3% of the area, which is >60% greater than that achieved by PHYSAT (Table 1). This large number of PFT retrievals is due to the SVD performed by PHYSTWO, which does not produce undetermined values as is the case with PHYSAT as an effect of removing the Chl-a through the division of the satellite nLw observations by the reference models (Equation (2)). However, PHYSTWO retrievals with unadjusted orthomodels are similar to PHYSAT estimations and, therefore, they still do not agree with the in-situ observations. The percentage of agreement in the case of Diatoms slightly increases with regard to PHYSAT but remains low (~19.7%), whereas there is no agreement with the retrievals for Nanoeukaryotes (Table 3). Only the *Prochlorococcus* agreement increased over 50% but, as observed in Figure 8, there were also an important number of retrievals close the coast in waters with high Chl-a concentration, where this group is unlikely to be found.

When adjusted orthomodels are used, PHYSTWO retrievals of Diatoms and Nanoeukaryotes contribute to the increase in the percentage of agreement to 81.6% and 68.9%, respectively. In the case of *Synechococcus*, the agreements remain similar, while those of *Prochlorococcus* decrease, although they are still higher than those reached by PHYSAT. Thus, PHYSTWO produces a PFT distribution closer to that expected from the in situ measurements during the PHYTOFRONT cruise and from the known environmental preferences of the PFTs (Figure 13). This type of PFT estimation displays a dominance of Diatoms followed by Nanoeukaryotes in coastal areas. The extension of coastal waters through the filament is also dominated by Diatoms with small patches of *Phaeocystis*. Finally, the most offshore areas are dominated by *Synechococcus* and *Prochlorococcus*, including the intrusions of oceanic waters in the southern section of the eddy, an effect of its circulation pattern [29].

4. Discussion

Since the first satellite-color missions, satellite data in the visible spectrum have been used to provide valuable information about the optical properties and the main compounds found in the surface ocean. Chl-a concentration as a proxy of phytoplankton biomass has been by far the most used product from ocean-color satellites and, as the first biological variable measured from space, it triggered an important breakthrough in our knowledge and perception of the ocean as the largest ecosystem on Earth. An evolving interest in retrieving information on other properties, including the composition of phytoplanktonic communities, has emerged in recent years [1]. This interest relates to the need of characterizing the spatial and temporal variability of primary producers, which in turn influence biogeochemical cycles, ecosystem structures, trophic transfers, and the climate.

Satellite Chl-a measurements are currently straightforward and obtained through empirical algorithms [43,44] since this is the main photosynthetic pigment in the upper ocean [36]. The Chl-a signal in the nLw radiances retrieved by satellite dominates most of the ocean landscape and it overshadows the signals of other phytoplanktonic (accessory) pigments, which are useful in PTF identification. The PHYSAT approach removes the effect of Chl-a by normalizing nLw data using a reference Chl-a spectrum between 412 and 555 nm as a way to highlight the signal of the accessory pigments [22]. However, the normalization process makes the information on the environmental Chl-a concentration unavailable for further PFT identification. Chl-a concentration is a very important piece of information on the environmental regime of the phytoplanktonic community under study since, in general, the dominance of a given PFT is associated to particular nutrients and Chl-a concentrations (that is, diatoms are likely the dominant PFT in high nutrient and Chl-a conditions while cyanobacteria are expected to dominate under low nutrient and Chl-a concentrations) [45,46].

In the present study, we have applied PHYSAT to an area ranging from nutrient-rich nearshore waters to offshore nutrient-poor waters in a temperate coastal upwelling region. We have shown that the main errors in the PHYSAT estimates are caused by PFT assignments to environments where these PFTs are not dominant or less likely to live. We have also provided evidence supporting the idea that the signal of the accessory pigments is spectrally uncorrelated with that of Chl-a so that both signals (Chl-a and non-Chl-a) conform independent modes of spectral variation in the nLw data.

Based on this, the proposed PHYSTWO approximation isolates the Chl-a signal into a dominant (first) spectral orthogonal mode (U_1) and together with a second mode (U_2), containing the signal of the accessory pigments, performs a PFT estimation in the U_1–U_2 plane. Thus, the main advantage of PHYSTWO over PHYSAT is that it does not require the removal of the Chl-a spectral effect from the nLw data. This allows PFT estimations over a wide scope of environmental conditions with different Chl-a concentration. In theory, the U_1–U_2 plane contains all possible spectral combinations of the seven nLw wavelengths, but it is also a bi-dimensional arrangement for all possible combinations of the spectral signal coming from different phytoplankton communities under the diverse ecological regiments that characterize the oceans. Thus, PHYSTWO appears to be more suitable than PHYSAT to perform PFT estimations in complex regions like coastal upwelling regions where the optical properties change quickly from high to low Chl-a concentrations in a coastal to offshore gradient.

The presence of other optically significant agents in surface waters, such as detritus and colored dissolved organic matter (CDOM) may bias the remote detection of phytoplankton groups [25]. In the case of the upwelling system off central Chile, our observations indicated an increase of CDOM with depth and had a low correlation with the satellite surface Chl-a. Despite this, the low CDOM levels in surface waters may account for the spectral maximum at 469 nm observed in the regional reference model nLw^{upw}, since the absorption by CDOM is large in the blue part of the spectrum. However, the PFT estimation using this regional model produces a similar estimation to that using the standard nLw reference model, which suggests that the CDOM represents only a low interference for PFT estimation in this region. It is likely, however, that this is not the case for all coastal waters, especially for those influenced by river discharges, where higher CDOM concentrations are expected and their optical characteristics could be more similar to the case-2 waters. Unfortunately, the few CDOM observations included in our study do not allow an appropriated analysis of the CDOM interference but we expect that further investigations will allow the inclusion of CDOM as an additional PHYSTWO orthomodel and, with it, improve the PFT estimations.

The PFT estimation through PHYSTWO relies on the use of orthomodels in the dimensionless plane U_1–U_2. This plane is a continuous field which provides information on the spatial changes associated with the transition between different spectral regions. As such, this approximation could be used to identify the relative abundance over a greater number of phytoplankton assemblages compared to the categorical dominance obtained by PHYSAT. Hence, PHYSTWO can potentially provide information on the spatial succession or co-dominance in the phytoplankton communities. However, the orthomodels used in this study to perform the PFT estimation are in effect an adjusted version of the PHYSAT acceptable ranges [22,24,27], so they cannot identify the co-dominance regions in their current state. Future PHYSTWO versions could be evolved towards the estimation of mixed PFT dominance in different assemblages when detailed information on in situ phytoplankton community structure, or on accessory pigments, becomes available to build new orthomodels.

The adjustments and validation of the PHYSTWO orthomodels in a given region rely on the availability of in-situ information on the phytoplankton community structure, which in general, is scarce for the oceans as a whole. In the present study, the orthomodel adjustments were done employing different sources of data. The *Phaeocystis* orthomodel was adjusted using the global databases collected in the last 13 years; those of Cyanobacteria orthomodels were adjusted using genomic studies performed at the basin scale; and in the case of Diatoms and Nanoeukaryotes, the adjustments were done using in vivo fluorescence obtained from a spectro-fluorometer (FluoroProbe profiler; bbe Moldaenke GmbH, Kiel, Germany). Altogether, the improvement of PHYSTWO over PHYSAT in PFT identification in the region of study and using local, as well as global databases, suggest that it has the potential for further improvements with larger datasets and that it can be applied to other regions of the oceans, even in its present format. On the one hand, our results indicated that spectro-fluorometers provide suitable information for this purpose because it reports the relative abundance of the phytoplankton groups in terms of their chlorophyll content. This is a desirable characteristic that allows a direct estimation of the degree of local spectral dominance without

the requirement of bio-volume transformations from biomass information. Previous studies on the comparison of different techniques for assessing phytoplankton community structures have shown that FluoroProbe is potentially capable of determining its general characteristics when compared to HPLC analyses [47–49]. However, its application improves greatly when species-specific calibrations under different environmental conditions are considered [49]. In the case of the FluoroProbe data used in this study, a comparison of the Chl-a estimates with those obtained from fluorometric analyses revealed very similar patterns of distribution, though the magnitudes were higher in the first case. Additionally, the surface satellite and in situ Chl-a data presented a similar distribution [29]. Differences in the magnitude of the estimates do not affect PFT assignation since orthomodels represent the relative abundances of the different groups with respect to the total.

The PHYSTWO approach incorporates a synthetic matrix, initially based on the PHYSAT acceptable ranges, which after it undergoes an SVD, produces orthomodels with reference values that are useful to perform PFT identification. These orthomodels operate in the U_1–U_2 plane and they can be easily adjusted in this plane using additional new in-situ observations and published data. Because the U_1–U_2 plane has a finite space, the changes in the scope of any orthomodel (that is, with the addition or removal of orthomodel points) affect the scope of the neighbor orthomodels, even though the latter is not modified. After adjustment, we have shown that the orthomodels are suitable to produce accurate PFT estimations in environments with high Chl-a concentrations. However, these adjustments do not constrain PFT identification to this kind of environment. The orthomodels were modified mostly in their top-left ends to ensure an adequate PFT identification in high Chl-a waters (>3 mg·m^{-3}), leaving their bottom-right values unaltered, where the reference values for clear waters (Chl-a concentrations between 0.1 and 3 mg·m^{-3}) are located. Thus, PFT identification by PHYSTWO and PHYSAT remained similar in environments with low Chl-a concentrations. Altogether, the adjustments performed to orthomodels make PHYSTWO suitable to operate over a wider range of environments than PHYSAT and it does not require the use of regional reference models to acknowledge the particular optical differences between ocean regions (Figure 13c). However, further adjustments to PHYSTWO orthomodels beyond those performed in the present study may be required to increase the improvement in PFT estimation. Potential users of PHYSTWO need to determine if the error inherent in this technique is acceptable for the type of research being conducted. Matlab codes for PHYSTWO and PHYSAT are freely-available at https://www.dropbox.com/sh/1j7sn3uo7fa60r5/AABwub8DEg1YOqsICY_PankEa?dl=0.

5. Conclusions

Knowledge on the phytoplankton community structure in the oceans and on its spatial and temporal scales of variability directly contribute to more accurate assessments of the state of pelagic ecosystems, the extent of carbon cycling, and the impacts of climate change. PHYSAT is a pioneering method in the identification of PFTs using remote sensing retrievals of radiance reflectance spectra. PHYSAT uses reference data from open ocean environments and it fails to produce appropriate estimates in waters that are optically more complex, such as those in highly productive coastal waters. An adaption of the PHYSAT method to more optically complex waters through the construction of regional reference models does not substantially improve PFT estimation in the case of coastal upwelling regions, as assessed during the present study. The proposed PHYSTWO method represents an alternative approach applicable to systems with a larger range of Chl-a concentration and without requiring the construction of additional regional reference models. Since the signal of Chl-a and non-chlorophyll pigments conform to different modes of the spectral variability retrieved by satellites, PHYSTWO separates these signal by employing an orthogonal decomposition of space–nLw matrices constructed with satellite nLw information. With the use of adjusted orthomodels in the dimensionless plane formed by the first two modes. PHYSTWO produces PFT retrievals with a better agreement regarding in-situ observations than those obtained by PHYSAT in the coastal upwelling region under study. Furthermore, PHYSTWO reduces the number of indeterminate assignments generated

by PHYSAT by more than 60%. PHYSTWO orthomodels can be easily adjusted to improve PFT determination and can be complemented by the incorporation of new PTFs beyond the six categories defined by PHYSAT if additional in situ data are employed. These aspects will require further assessments of the PHYSTWO method in different regions and environmental conditions.

Supplementary Materials: The PHYSTWO codes are available online at http://www.mdpi.com/2072-4292/10/4/498/s1. Content: modis_phystwo.m and modis_physat.m contain Matlab codes for PHYSTWO and PHYSAT, respectively; PHYSTWO_synthetic_matrix (.mat and _adjusted.mat) contains orthomodel matrices; and text.m contains a demonstration script.

Acknowledgments: This work was supported by the National Fund for Scientific and Technological Development of Chile (FONDECYT projects 11130463 and 1151299) and IMO-Chile; the Coastal and Marine Research Institute (INVEMAR) and the Administrative Department of Science, Technology and Innovation (COLCIENCIAS) of Colombia (project COLCIENCIAS-INVEMAR FP44842-499-2014), with funds from the Fondo Nacional de Financiamiento para la Ciencia, la Tecnología y la Innovación Francisco José de Caldas. The authors thanks the NASA Goddard Space Flight Center, the Ocean Ecology Laboratory and Ocean Biology Processing Group for the production and distribution of the Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Rrs L2 data; NASA OB.DAAC, Greenbelt, MD, USA. DOI: 10.5067/AQUA/MODIS_OC.2014.0.

Author Contributions: M.C.-R. and R.L. conceived and designed the numerical experiments and analysis; M.C.-R. performed the analysis; C.E.M. and S.H. contributed with in situ data; V.A. performed laboratory analysis; and M.C.-R., C.M. and R.L. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest

Appendix A. The *R* Matrix

PHYSTWO is based on an EOF decomposition of a two-dimensional space–nLw matrix *R*. This matrix is built by reordering the satellite nLw fields (longitude–latitude matrices) at 412, 443, 469, 488, 531, 547, and 555 nm wavelengths:

$$R(\lambda)^p = \begin{vmatrix} nLw(412)^1 & nLw(443)^1 & nLw(469)^1 & nLw(488)^1 & nLw(531)^1 & nLw(547)^1 & nLw(555)^1 \\ nLw(412)^2 & nLw(443)^2 & nLw(469)^2 & nLw(488)^2 & nLw(531)^2 & nLw(547)^2 & nLw(555)^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ nLw(412)^p & nLw(443)^p & nLw(469)^p & nLw(488)^p & nLw(531)^p & nLw(547)^p & nLw(555)^p \end{vmatrix} \quad \text{(A1)}$$

In the *R* matrix, each row contains the nLw data for each position (or pixel) and *P* corresponds to the total number of positions in the region. The rows were excluded from the *R* matrix in the positions where nLw (555) was >1.3 $W \cdot m^{-2} \cdot \mu m^{-1} \cdot sr^{-1}$ and AOT was >0.15 to reduce the presence of biased values arising from high concentrations of suspended sediments or by errors in the atmospheric correction, as recommended for the PHYSAT method [25,27]. To avoid numerical dominance caused by the high values, each column (or wavelength) of the *R* matrix was standardized previously to the orthogonal decomposition by removing its corresponding mean (μ (λ)) and dividing it by the standard deviation (σ (λ)) of each wavelength, obtaining the matrix *N*:

$$N(\lambda)^P = \frac{\left(R(\lambda)^P - \mu(\lambda)\right)}{\sigma(\lambda)}. \quad \text{(A2)}$$

Appendix B. The Unadjusted Synthetic Matix

For each PFT, a typical anomaly (PFT_{ta}, where *PFT* can be *NEU*, *PRO*, *SLC*, *DIA*, *PHA*, or *COB*) is calculated by averaging the upper and lower limits of the acceptable ranges of *Ra* at each wavelength (continuous bold lines in Figure 4). Assuming that the Chl-a concentration is an indicator of the contribution of all groups that compose a phytoplankton community, it is possible to obtain a synthetic matrix (Rs) containing the typical nLw values of each PFT at several Chl-a concentrations by multiplying PFT_{ta} by the Chl-a spectrum represented in the nLw^{ref}. Here, we considered an nLw^{ref} with 300 Chl-a spectra at concentrations in the range of the GeP&CO measurements (between 0.01 and

3 mg·m^{-3}, c = 0.01, 0.02..., 3 mg·m^{-3}). Thus, *Rs* results in a matrix of 1800 rows (300 for each of the six PFTs) by seven columns (wavelength):

$$Rs(\lambda)^c = \begin{vmatrix} NEU_{ta}(412) * nLw^{ref,c=0.01} & NEU_{ta}(443) * nLw^{ref,c=0.01} & \cdots & NEU_{ta}(555) * nLw^{ref,c=0.01} \\ NEU_{ta}(412) * nLw^{ref,c=0.02} & NEU_{ta}(443) * nLw^{ref,c=0.02} & \cdots & NEU_{ta}(555) * nLw^{ref,c=0.02} \\ \vdots & \vdots & & \vdots \\ NEU_{ta}(412) * nLw^{ref,c=3} & NEU_{ta}(443) * nLw^{ref,c=3} & \cdots & NEU_{ta}(555) * nLw^{ref,c=3} \\ PRO_{ta}(412) * nLw^{ref,c=0.01} & PRO_{ta}(443) * nLw^{ref,c=0.01} & \cdots & PRO_{ta}(555) * nLw^{ref,c=0.01} \\ \vdots & \vdots & & \vdots \\ PRO_{ta}(412) * nLw^{ref,c=3} & PRO_{ta}(443) * nLw^{ref,c=3} & \cdots & PRO_{ta}(555) * nLw^{ref,c=3} \\ SLC_{ta}(412) * nLw^{ref,c=0.01} & SLC_{ta}(443) * nLw^{ref,c=0.01} & \cdots & SLC_{ta}(555) * nLw^{ref,c=0.01} \\ \vdots & \vdots & & \vdots \\ SLC_{ta}(412) * nLw^{ref,c=3} & SLC_{ta}(443) * nLw^{ref,c=3} & \cdots & SLC_{ta}(555) * nLw^{ref,c=3} \\ DIA_{ta}(412) * nLw^{ref,c=0.01} & DIA_{ta}(443) * nLw^{ref,c=0.01} & \cdots & DIA_{ta}(555) * nLw^{ref,c=0.01} \\ \vdots & \vdots & & \vdots \\ DIA_{ta}(412) * nLw^{ref,c=3} & DIA_{ta}(443) * nLw^{ref,c=3} & \cdots & DIA_{ta}(555) * nLw^{ref,c=3} \\ PHA_{ta}(412) * nLw^{ref,c=0.01} & PHA_{ta}(443) * nLw^{ref,c=0.01} & \cdots & PHA_{ta}(555) * nLw^{ref,c=0.01} \\ \vdots & \vdots & & \vdots \\ PHA_{ta}(412) * nLw^{ref,c=3} & PHA_{ta}(443) * nLw^{ref,c=3} & \cdots & PHA_{ta}(555) * nLw^{ref,c=3} \\ COB_{ta}(412) * nLw^{ref,c=0.01} & COB_{ta}(443) * nLw^{ref,c=0.01} & \cdots & COB_{ta}(555) * nLw^{ref,c=0.01} \\ \vdots & \vdots & & \vdots \\ COB_{ta}(412) * nLw^{ref,c=3} & COB_{ta}(443) * nLw^{ref,c=3} & \cdots & COB_{ta}(555) * nLw^{ref,c=3} \end{vmatrix} \quad \text{(A3)}$$

The resulting *Rs* matrix is a concentration–wavelength matrix whose values have the same units (W·m^{-2}·µm^{-1}·sr^{-1}) of the *R* matrix (Figure A1).

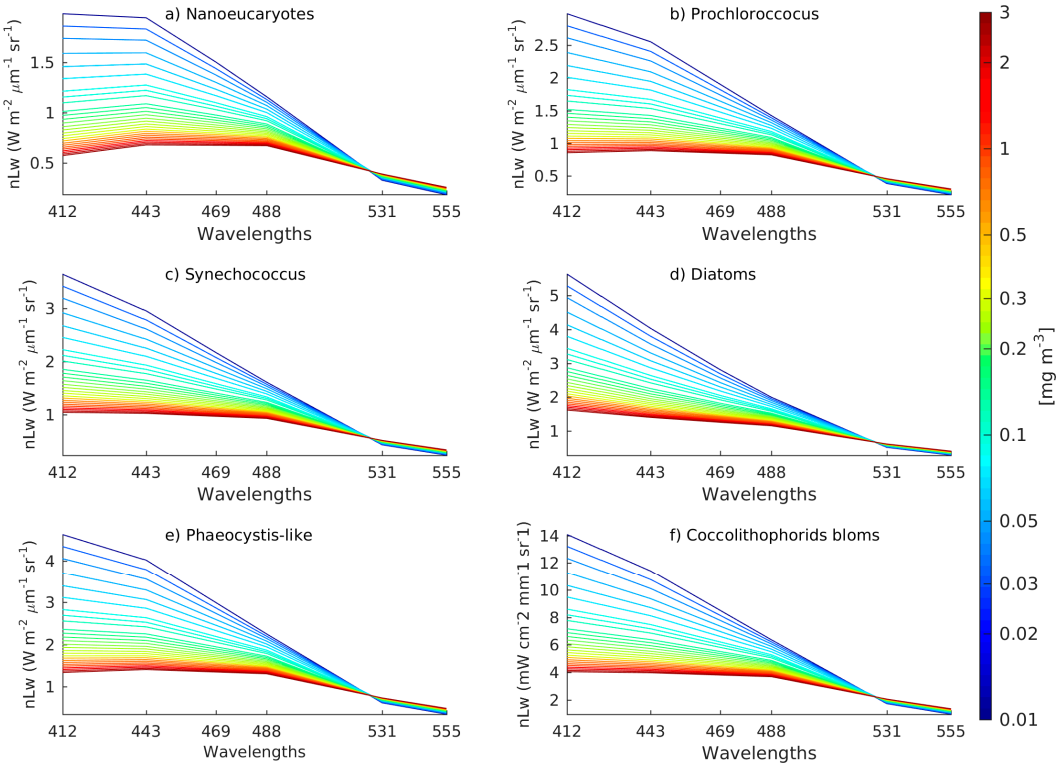

Figure A1. The typical nLw radiances for PFTs in environments with Chl-a concentrations in the range between 0.01 and 3 mg·m^{-3}, contained in the *Rs* synthetic matrix.

References

1. IOCCG. *Phytoplankton Functional Types from Space*; Sathyendranath, S., Ed.; Reports of the International Ocean-Colour Coordinating Group; IOCCG: Dartmouth, NS, Canada, 2014; Volume 15.
2. 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. [CrossRef]
3. Quéré, C.L.; Harrison, S.P.; Colin Prentice, I.; Buitenhuis, E.T.; Aumont, O.; Bopp, L.; Claustre, H.; Cotrim Da Cunha, L.; Geider, R.; Giraud, X.; et al. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* **2005**, *11*, 2016–2040. [CrossRef]
4. Litchman, E.; Klausmeier, C.A. Trait-Based Community Ecology of Phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* **2008**, *39*, 615–639. [CrossRef]
5. Sieburth, J.M.; Smetacek, V.; Lenz, J. Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnol. Oceanogr.* **1978**, *23*, 1256–1263. [CrossRef]
6. Finkel, Z.V.; Beardall, J.; Flynn, K.J.; Quigg, A.; Rees, T.A.V.; Raven, J.A. Phytoplankton in a changing world: Cell size and elemental stoichiometry. *J. Plankton Res.* **2010**, *32*, 119–137. [CrossRef]
7. Follows, M.J.; Dutkiewicz, S. Modeling Diverse Communities of Marine Microbes. *Annu. Rev. Mar. Sci.* **2011**, *3*, 427–451. [CrossRef] [PubMed]
8. Nair, A.; Sathyendranath, S.; Platt, T.; Morales, J.; Stuart, V.; Forget, M.-H.; Devred, E.; Bouman, H. Remote sensing of phytoplankton functional types. *Remote Sens. Environ.* **2008**, *112*, 3366–3375. [CrossRef]
9. Aiken, J.; Pradhan, Y.; Barlow, R.; Lavender, S.; Poulton, A.; Holligan, P.; Hardman-Mountford, N. Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995–2005. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2009**, *56*, 899–917. [CrossRef]
10. Mouw, C.B.; Hardman-Mountford, N.J.; Alvain, S.; Bracher, A.; Brewin, R.J.W.; Bricaud, A.; Ciotti, A.M.; Devred, E.; Fujiwara, A.; Hirata, T.; et al. A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. *Front. Mar. Sci.* **2017**, *4*. [CrossRef]
11. Bracher, A.; Bouman, H.A.; Brewin, R.J.W.; Bricaud, A.; Brotas, V.; Ciotti, A.M.; Clementson, L.; Devred, E.; Di Cicco, A.; Dutkiewicz, S.; et al. Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development. *Front. Mar. Sci.* **2017**, *4*. [CrossRef]
12. Chase, A.; Boss, E.; Zaneveld, R.; Bricaud, A.; Claustre, H.; Ras, J.; Dall'Olmo, G.; Westberry, T.K. Decomposition of in situ particulate absorption spectra. *Methods Oceanogr.* **2013**, *7*, 110–124. [CrossRef]
13. 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. [CrossRef] [PubMed]
14. Uitz, J.; Stramski, D.; Reynolds, R.A.; Dubranna, J. Assessing phytoplankton community composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-sensing reflectance in open-ocean environments. *Remote Sens. Environ.* **2015**, *171*, 58–74. [CrossRef]
15. Di Cicco, A.; Sammartino, M.; Marullo, S.; Santoleri, R. Regional Empirical Algorithms for an Improved Identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea Using Satellite Data. *Front. Mar. Sci.* **2017**, *4*. [CrossRef]
16. Sammartino, M.; Di Cicco, A.; Marullo, S.; Santoleri, R. Spatio-temporal variability of micro-, nano- and pico-phytoplankton in the Mediterranean Sea from satellite ocean colour data of SeaWiFS. *Ocean Sci. Discuss.* **2015**, *12*, 161–201. [CrossRef]
17. Brotas, V.; Brewin, R.J.W.; Sá, C.; Brito, A.C.; Silva, A.; Mendes, C.R.; Diniz, T.; Kaufmann, M.; Tarran, G.; Groom, S.B.; et al. Deriving phytoplankton size classes from satellite data: Validation along a trophic gradient in the eastern Atlantic Ocean. *Remote Sens. Environ.* **2013**, *134*, 66–77. [CrossRef]
18. Brewin, R.J.W.; Sathyendranath, S.; Lange, P.K.; Tilstone, G. Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep Sea Res. Part Oceanogr. Res. Pap.* **2014**, *85*, 72–79. [CrossRef]
19. Laney, S.R.; Sosik, H.M. Phytoplankton assemblage structure in and around a massive under-ice bloom in the Chukchi Sea. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2014**, *105*, 30–41. [CrossRef]
20. Brewin, R.J.W.; Hardman-Mountford, N.J.; Lavender, S.J.; Raitsos, D.E.; Hirata, T.; Uitz, J.; Devred, E.; Bricaud, A.; Ciotti, A.; Gentili, B. An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* **2011**, *115*, 325–339. [CrossRef]

21. Kostadinov, T.S.; Cabré, A.; Vedantham, H.; Marinov, I.; Bracher, A.; Brewin, R.J.W.; Bricaud, A.; Hirata, T.; Hirawake, T.; Hardman-Mountford, N.J.; et al. Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and Earth System Models. *Remote Sens. Environ.* **2017**, *190*, 162–177. [CrossRef]
22. Alvain, S.; Moulin, C.; Dandonneau, Y.; Bréon, F.M. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. Part Oceanogr. Res. Pap.* **2005**, *52*, 1989–2004. [CrossRef]
23. Dandonneau, Y.; Deschamps, P.-Y.; Nicolas, J.-M.; Loisel, H.; Blanchot, J.; Montel, Y.; Thieuleux, F.; Becu, G. Seasonal and interannual variability of ocean color and composition of phytoplankton communities in the North Atlantic, equatorial Pacific and South Pacific. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2004**, *51*, 303–318. [CrossRef]
24. Alvain, S.; Moulin, C.; Dandonneau, Y.; Loisel, H. Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Glob. Biogeochem. Cycles* **2008**, *22*. [CrossRef]
25. 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. [CrossRef] [PubMed]
26. Siegel, D.A.; Maritorena, S.; Nelson, N.B.; Behrenfeld, M.J. Independence and interdependencies among global ocean color properties: Reassessing the bio-optical assumption. *J. Geophys. Res. Oceans* **2005**, *110*, C07011. [CrossRef]
27. Navarro, G.; Alvain, S.; Vantrepotte, V.; Huertas, I.E. Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. *Remote Sens. Environ.* **2014**, *152*, 557–575. [CrossRef]
28. Catherine, A.; Escoffier, N.; Belhocine, A.; Nasri, A.B.; Hamlaoui, S.; Yéprémian, C.; Bernard, C.; Troussellier, M. On the use of the FluoroProbe®, a phytoplankton quantification method based on fluorescence excitation spectra for large-scale surveys of lakes and reservoirs. *Water Res.* **2012**, *46*, 1771–1784. [CrossRef] [PubMed]
29. Morales, C.E.; Anabalón, V.; Bento, J.P.; Hormazabal, S.; Cornejo, M.; Correa-Ramírez, M.A.; Silva, N. Front-Eddy Influence on Water Column Properties, Phytoplankton Community Structure, and Cross-Shelf Exchange of Diatom Taxa in the Shelf-Slope Area Off Concepción (~36–37°S). *J. Geophys. Res. Oceans* **2017**. [CrossRef]
30. Böttjer, D.; Morales, C.E. Nanoplanktonic assemblages in the upwelling area off Concepción (∼36°S), central Chile: Abundance, biomass, and grazing potential during the annual cycle. *Prog. Oceanogr.* **2007**, *75*, 415–434. [CrossRef]
31. Collado-Fabbri, S.; Vaulot, D.; Ulloa, O. Structure and seasonal dynamics of the eukaryotic picophytoplankton community in a wind-driven coastal upwelling ecosystem. *Limnol. Oceanogr.* **2011**, *56*, 2334–2346. [CrossRef]
32. Morales, C.E.; Anabalón, V. Phytoplankton biomass and microbial abundances during the spring upwelling season in the coastal area off Concepción, central-southern Chile: Variability around a time series station. *Prog. Oceanogr.* **2012**, *92–95*, 81–91. [CrossRef]
33. Anabalón, V.; Morales, C.E.; González, H.E.; Menschel, E.; Schneider, W.; Hormazabal, S.; Valencia, L.; Escribano, R. Micro-phytoplankton community structure in the coastal upwelling zone off Concepción (central Chile): Annual and inter-annual fluctuations in a highly dynamic environment. *Prog. Oceanogr.* **2016**, *149*, 174–188. [CrossRef]
34. Vogt, M.; O'Brien, C.; Peloquin, J.; Schoemann, V.; Breton, E.; Estrada, M.; Gibson, J.; Karentz, D.; Van Leeuwe, M.A.; Stefels, J.; et al. Global marine plankton functional type biomass distributions: *Phaeocystis* spp. *Earth Syst. Sci. Data* **2012**, *4*, 107–120. [CrossRef]
35. Peperzak, L.; van der Woerd, H.J.; Timmermans, K.R. Disparities between in situ and optically derived carbon biomass and growth rates of the prymnesiophyte *Phaeocystis globosa*. *Biogeosciences* **2015**, *12*, 1659–1670. [CrossRef]
36. Gordon, H.R.; Clark, D.K.; Brown, J.W.; Brown, O.B.; Evans, R.H.; Broenkow, W.W. Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates. *Appl. Opt.* **1983**, *22*, 20. [CrossRef] [PubMed]
37. Morel, A. Optical Modeling of the Upper Ocean in Relation to Its Biogenous. Matter Content (Case I Waters. *J. Geophys. Res.* **1988**, *93*, 10749–10768. [CrossRef]
38. Bjornsson, H.; Venegas, S. *A Manual for EOF and SVD Analyses of Climate Data*; C2GCR Report Ser.; McGill University: Montreal, QC, Canada, 1997; Volume 97–1, pp. 1–52.

39. Hutchins, D.A.; Hare, C.E.; Weaver, R.S.; Zhang, Y.; Firme, G.F.; DiTullio, G.R.; Alm, M.B.; Riseman, S.F.; Maucher, J.M.; Geesey, M.E.; et al. Phytoplankton iron limitation in the Humboldt Current and Peru Upwelling. *Limnol. Oceanogr.* **2002**, *47*, 997–1011. [CrossRef]
40. Schoemann, V.; Becquevort, S.; Stefels, J.; Rousseau, V.; Lancelot, C. *Phaeocystis* blooms in the global ocean and their controlling mechanisms: A review. *J. Sea Res.* **2005**, *53*, 43–66. [CrossRef]
41. Heywood, J.L.; Zubkov, M.V.; Tarran, G.A.; Fuchs, B.M.; Holligan, P.M. Prokaryoplankton standing stocks in oligotrophic gyre and equatorial provinces of the Atlantic Ocean: Evaluation of inter-annual variability. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2006**, *53*, 1530–1547. [CrossRef]
42. Bibby, T.S.; Zhang, Y.; Chen, M. Biogeography of Photosynthetic Light-Harvesting Genes in Marine Phytoplankton. *PLoS ONE* **2009**, *4*, e4601. [CrossRef] [PubMed]
43. 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. Oceans* **1998**, *103*, 24937–24953. [CrossRef]
44. Hu, C.; Lee, Z.; Franz, B. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference: A novel ocean chlorophyll-a algorithm. *J. Geophys. Res. Oceans* **2012**, *117*. [CrossRef]
45. 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. Model.* **2010**, *221*, 1472–1483. [CrossRef]
46. Sathyendranath, S.; Cota, G.; Stuart, V.; Maass, H.; Platt, T. Remote sensing of phytoplankton pigments: A comparison of empirical and theoretical approaches. *Int. J. Remote Sens.* **2001**, *22*, 249–273. [CrossRef]
47. Liu, X.; Huang, B.; Liu, Z.; Wang, L.; Wei, H.; Li, C.; Huang, Q. High-resolution phytoplankton diel variations in the summer stratified central Yellow Sea. *J. Oceanogr.* **2012**, *68*, 913–927. [CrossRef]
48. See, J.H.; Campbell, L.; Richardson, T.L.; Pinckney, J.L.; Shen, R.; Guinasso, N.L. Combining new technologies for determination of phytoplankton community structure in the northern Gulf of Mexico 1: Phytoplankton community structure. *J. Phycol.* **2005**, *41*, 305–310. [CrossRef]
49. Escoffier, N.; Bernard, C.; Hamlaoui, S.; Groleau, A.; Catherine, A. Quantifying phytoplankton communities using spectral fluorescence: The effects of species composition and physiological state. *J. Plankton Res.* **2015**, *37*, 233–247. [CrossRef]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).