



Article

Evaluation of Satellite-Based Algorithms to Retrieve Chlorophyll-a Concentration in the Canadian Atlantic and Pacific Oceans

Stephanie Clay ^{1,†}, Angelica Peña ², Brendan DeTracey ³ and Emmanuel Devred ^{3,*,†}

- Takuvik Joint International Laboratory, Laval University-CNRS, Quebec City, QC G1V 0A6, Canada; Stephanie.Clay.1@ulaval.ca
- Institute of Ocean Sciences, Fisheries and Oceans Canada, Sidney, BC V8L 5T5, Canada; Angelica.Pena@dfo-mpo.gc.ca
- Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS B2Y 4A2, Canada; Brendan.DeTracey@dfo-mpo.gc.ca
- * Correspondence: emmanuel.devred@dfo-mpo.gc.ca; Tel.: +1-902-426-4681
- † These authors contributed equally to this work.

Received: 3 October 2019; Accepted: 1 November 2019; Published: 7 November 2019



Abstract: Remote-sensing reflectance data collected by ocean colour satellites are processed using bio-optical algorithms to retrieve biogeochemical properties of the ocean. One such important property is the concentration of chlorophyll-a, an indicator of phytoplankton biomass that serves a multitude of purposes in various ocean science studies. Here, the performance of two generic chlorophyll-a algorithms (i.e., a band ratio one, Ocean Colour X (OCx), and a semi-analytical one, Garver-Siegel Maritorena (GSM)) was assessed against two large in situ datasets of chlorophyll-a concentration collected between 1999 and 2016 in the Northeast Pacific (NEP) and Northwest Atlantic (NWA) for three ocean colour sensors: Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), and Visible Infrared Imaging Radiometer Suite (VIIRS). In addition, new regionally-tuned versions of these two algorithms are presented, which reduced the mean error (mg m⁻³) of chlorophyll-a concentration modelled by OCx in the NWA from -0.40, -0.58 and -0.45 to 0.037, -0.087 and -0.018 for MODIS, SeaWiFS, and VIIRS respectively, and -0.34 and -0.36 to -0.0055 and -0.17 for SeaWiFS and VIIRS in the NEP. An analysis of the uncertainties in chlorophyll-a concentration retrieval showed a strong seasonal pattern in the NWA, which could be attributed to changes in phytoplankton community composition, but no long-term trends were found for all sensors and regions. It was also found that removing the 443 nm waveband for the OCx algorithms significantly improved the results in the NWA. Overall, GSM performed better than the OCx algorithms in both regions for all three sensors but generated fewer chlorophyll-a retrievals than the OCx algorithms.

Keywords: ocean colour; satellite-derived chlorophyll-a concentration; algorithm evaluation; Northwest Atlantic; Northeast Pacific

1. Introduction

The product derived from satellite ocean colour that is the most used is undoubtedly chlorophyll-a (*chla*) concentration, an index of phytoplankton biomass, which has numerous applications in biogeochemical oceanography, such as phytoplankton ecology and phenology [1,2], carbon cycles [3], climate change, transfer of energy to higher trophic levels, and water quality [4]. Satellite-derived *chla* concentration is a mature product of ocean colour that is used not only by experts in the field of bio-optics but also by the entire oceanographic community in various ways, such as modelling

Remote Sens. 2019, 11, 2609 2 of 29

and data assimilation [5], fisheries applications [6], and ecosystem management [7]. For instance, European programs such as Copernicus (https://www.copernicus.eu/) or NOAA's COASTWATCH (https://coastwatch.noaa.gov/cw/index.html) provide daily *chla* data that are accessible to the public for fisheries and water quality applications. In Canada, Fisheries and Oceans have been relying on ocean colour for several decades to monitor the state of the marine ecosystem [8,9], and in particular on *chla* concentration, and its derived product primary production, to infer ecosystem indicators that are at the foundation of ocean management.

Chla concentration is derived from remote sensing reflectance (R_{rs}) , which is the ratio of water-leaving radiance to downwelling irradiance corrected for sun geometry to allow for comparisons independent of locations, times and dates. Several approaches to infer chla from R_{rs} have been developed and embedded by space agencies in their data processing software, including band ratio [10] and semi-analytical models [11]. As its name suggests, band ratio algorithms exploit the ratio of wavebands in the blue and green to retrieve chla; the Ocean Colour X (OCx) (x stands for 2, 3, or 4 and indicates the number of bands that were used in the algorithm) suite of empirical algorithms have been developed by the National Aeronautics and Space Administration (NASA) using a global dataset of in situ measurements of chla concentration fitted to remote sensing reflectance (NOMAD, the NASA bio-Optical Marine Algorithm Dataset [12]) and a fourth-degree polynomial expression. On the other hand, semi-analytical algorithms (e.g., Garver-Siegel Maritorena (GSM)) consist of optimizing bio-optical parameters (including chla concentration) in an approximate solution of the radiative transfer equation to match modelled reflectance to the reflectance measured by the satellite. This type of algorithm has the advantage of decoupling the contribution of the optically active components (i.e., phytoplankton, non-algal particles and coloured dissolved organic carbon) such that chla concentration should, in theory, be retrieved with higher accuracy than the band ratio algorithms. These two types of approaches are well-suited for the so-called case-1 waters (i.e., waters where chla concentration drives the optical characteristics of bulk seawater) but do not perform as well in case-2 waters (i.e., where optical signals of marine components are uncorrelated). For case-2 waters, models that use fluorescence [13] or more advanced statistical methods, such as neural networks [14] or principal component analysis [15], have been demonstrated to perform better than band ratio or semi-analytical approaches. Note that the performance of algorithms that use remote sensing reflectance will be inherently dependent on the performance of the atmospheric correction procedure, which will not be addressed here.

It has been shown that OCx [15,16] and GSM [17] algorithms contain regional biases, such that even if its overall performance is satisfying, application to a given region results in systematic bias. Here, we assessed the performance of the OCx and GSM algorithms in Canadian waters (Northwest Atlantic (NWA) and Northeast Pacific (NEP)) using a dataset of *in situ chla* concentration that was collected by Fisheries and Oceans Canada as part of their monitoring of the marine ecosystem. Performance of these algorithms was evaluated for NASA's three sensors, namely the Sea-viewing Wide-field-of-View Sensor (SeaWiFS, 1998–2010), the Moderate Resolution Imaging Spectrometer on the Aqua platform (MODIS-Aqua, 2002–current) and the Visible and Infrared Imaging Spectrometer on the NPP platform (VIIRS, 2012–current). Uncertainties of the algorithms with respect to environmental variables were also discussed.

2. Material and Methods

2.1. Regions of Interest

Fisheries and Oceans Canada (DFO) carries out sea-going expeditions as part of its oceanographic monitoring programs to obtain information on the health of the marine ecosystem to support decision-making in the management of the ocean. In the Northwest Atlantic (NWA) Ocean, more than a hundred stations are visited twice a year in spring and fall on the Scotian Shelf and once a year (in spring) in the Labrador Sea as part of the fieldwork of the Atlantic Zone Monitoring

Remote Sens. 2019, 11, 2609 3 of 29

Program (AZMP) and Atlantic Zone Offshore Monitoring Program (AZOMP) respectively (Figure 1). During these missions, a large number of physical, chemical and biological oceanographic parameters are collected [8]. As part of its remote sensing operations, DFO archives satellite data of ocean colour and sea-surface temperature in a region bounded by 39° to 82° north and 95° to 42° west. These data were used to assess the performance of ocean colour *chla* data products in the NWA. In a similar fashion, DFO runs monitoring programs in the NEP along Line P three times a year (in winter, spring and summer), off the west coast of Vancouver Island twice a year (in spring and summer) and in the Strait of Georgia three times a year (in spring, summer and fall), where physical, chemical and biological data are collected [18]. The in situ data from this region used in this study are bounded by 47° to 57° north, 148° to 123° west.

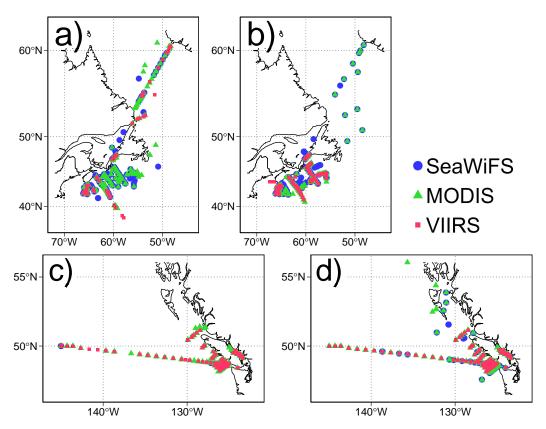


Figure 1. Locations of satellite matchups (a) in the Northwest Atlantic (NWA) from January to June, (b) from July to December, (c) in the Northeast Pacific (NEP) from January to June, and (d) from July to December.

2.2. In Situ Samples

DFO collects data on *chla* concentration using two methods, namely, Turner fluorescence (TF) and high-performance liquid chromatography (HPLC). While the most comprehensive dataset is the one obtained using the TF method (tens of thousands of data), here we used the dataset containing HPLC-derived *chla* concentration since it provides the most accurate estimates of *chla* and it is also the method recommended by space agencies for validation activities. A large dataset was compiled for the purpose of this study, consisting of the *in situ* samples gathered in the NWA and NEP which were analyzed in their respective DFO regional laboratories. In brief, water samples were rapidly filtered after collection through 25 mm GF/F glass fiber filters, immediately flash-frozen using liquid nitrogen, and stored in a $-80\,^{\circ}$ C freezer until analysis in the laboratory. The NWA samples were analyzed using a Beckman–Coulter Gold HPLC system (1998–2013) and more recently an Agilent 1200 (2013–2014) HPLC instrument following the protocol of Head and Horne [19] as modified by Stuart and Head [20]. Similarly, the NEP samples were analyzed using a Waters Alliance System HPLC

Remote Sens. 2019, 11, 2609 4 of 29

following the protocol of Zapata et al. [21]. The analysis was performed for 23 pigments for both locations, but only *chla* concentrations from both regions and fucoxanthin (fucox) concentrations from the NWA were used in this study. Note that chlorophyll-a concentration includes chlorophyll-a epimer and allomer pigments.

The complete dataset consisted of 1857 samples from the NWA between 1999 and 2014, and 1231 samples from the NEP between 2006 and 2016, all collected within ten metres from the surface. Both datasets showed a skewed normal distribution with a peak around 0.43 (0.42) mg m⁻³ in the NWA (NEP) Ocean and a similar range of variation of *chla* from 0.03 to 29.41 mg m⁻³, but samples collected in the NEP had a mean concentration (2.26 mg m⁻³) greater than in the NWA, where the mean concentration was 1.68 mg m⁻³ (Figure 2a). In the NWA, samples were collected mainly in spring (April to June) and fall (September to November) with a few exceptions of data collected in July and August, whereas the NEP dataset consisted of samples collected in winter (February to March), late spring-early summer (June to July) and late summer to early fall (August through October) (Figure 2b). Both datasets captured the seasonal cycle of phytoplankton, which was consistent with the wide range of biomass concentrations. The dataset was reduced to measurements collected within a day of satellite passes and ten kilometres from a valid pixel.

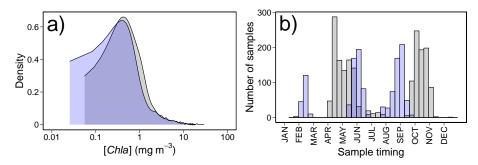


Figure 2. (a) Density function distribution of *in situ* chlorophyll-a concentration for the Northwest Atlantic (grey) and Northeast Pacific (lavender), and (b) total number of *in situ* samples available in the entire dataset for each month.

2.3. Satellite Matchups

The 2014 reprocessed datasets of SeaWiFS merged local area coverage (MLAC), MODIS-Aqua local area coverage (LAC), and VIIRS-Suomi NPP level-2 satellite scenes were downloaded from NASA's ocean colour L1/L2 browser (https://oceancolor.gsfc.nasa.gov/cgi/browse.pl?sen=am). For each image, a mask was applied to remove pixels with invalid data, defined as any pixel where one or more of the following criteria were met:

- 1. The pixel was marked by any of the following level 2 flags: atmospheric correction failure (ATMFAIL), deleted overlapping pixels (BOWTIEDEL), pixel overland (LAND), high sun glint (HIGLINT), pixel contains cloud or ice (CLDICE), radiance too high (HILT), high solar zenith (HISOLZEN), or high sensor zenith angle (HISATZEN).
- 2. The pixel contained a no-data (NA) value.
- 3. The pixel had more than one negative remote sensing reflectance (R_{rs}) value, implying the data might be flawed by the atmospheric correction procedure.

A search was performed for potential matches in any satellite image taken within a day of the $in \, situ$ sample, as exact sampling times were not always available in the database. In order to match a satellite value to an $in \, situ$ measurement, the satellite image was projected using a Gnomonic projection centred on the $in \, situ$ point, which stretched great circles to straight lines. In other words, the image is projected onto a plane tangent to the $in \, situ$ point, from which Cartesian coordinates (x,y) can be extracted for each corresponding (longitude, latitude) pair. For the sake of speed and simplicity, the Pythagorean theorem was then used to compute the distance between the $in \, situ$ point and each

Remote Sens. 2019, 11, 2609 5 of 29

satellite pixel. Though this method did not account for Earth's curvature, the computed distances at the maximum allowed radius around the *in situ* point (i.e., 10 km) were found to be within $\pm 23 \text{ m}$ of the more accurate geodetic distance as calculated by the distGeo() function using R Statistical Software [22]. A successful match was defined as the closest pixel, within a 10 km radius, that contained at least three valid pixels in a 3×3 pixel box extracted around the matching point. The distance between the *in situ* sample and the closest satellite pixel was recorded in the database. The median of the 3×3 pixel matrix was computed for each of the extracted remote sensing reflectance (R_{rs}) wavelengths, and the resulting dataset was considered as the total number of matches, mapped in Figure 1.

From the initial dataset of 1857 and 1231 samples for the NWA and NEP respectively, only 416 (SeaWiFS), 530 (MODIS) and 176 (VIIRS) *in situ* samples were matched to valid satellite pixels in the NWA, and 45 (SeaWiFS), 487 (MODIS) and 342 (VIIRS) in the NEP (Table 1). The small number of matchups compared to the initial dataset attests to the difficulty of compiling archives of satellite data that match *in situ* samples for validation purposes and emphasizes the need for recurrent measurements such as monitoring programs and moored automatic sensors.

Table 1. Specifications of sensors and matchup datasets. Bold and underlined numbers correspond to blue (λ_{blue}) and green (λ_{green}) bands used in the band ratio Ocean Colour X (OCx) algorithms respectively (see Section 2.4.1).

Sensor	Resolution (km)	Wavebands (nm)	Period (Year)	N	Period (Year)	N
			NWA		NEP	
MODIS	1.1	412, 443, 469, 488, 531, <u>547</u> , 555, 645, 667, 678	2002-2014	530	2006–2016	487
SeaWiFS	1.0	412, 443 , 490 , 510 , <u>555</u> , 670	1999-2010	416	2006-2010	45
VIIRS	0.75	410, 443 , 486 , <u>551</u> , 671	2012-2014	176	2012-2016	342

Finally, *chla* concentration was derived using the OCx algorithm (Section 2.4.1) and pixels with a coefficient of variation (computed as the ratio of the standard deviation to the mean of all pixels available in the 3×3 matrix) exceeding 0.5 were removed from the matchup dataset.

2.4. Chlorophyll-a Algorithms

Two main algorithms were tested: The Ocean Colour (OCx) [10] and the Garver–Siegel Maritorena (GSM) [23] algorithms, as well as variations of these original versions that were regionally-tuned. The variations consist of four new versions for OCx, following the same polynomial format in degrees 1–4 with regional coefficients, and three optimal versions of GSM, with different combinations of three modifications to the algorithm, described in Section 2.4.4. For the sake of simplicity, results from only six algorithms were analyzed in this study: the two original algorithms (OCx and GSM), two regionally-tuned counterparts for OCx, and two for GSM. The results of the remaining algorithms have been included in Appendix A. The algorithms were implemented using version 3.3.2 of R Statistical Software [24] and several functions and packages referenced in the text.

2.4.1. OCx-Type Algorithms

These empirical algorithms use a polynomial equation to express the log-10 transformed *chla* as a function of the log10-transformed ratio of blue to green remote sensing reflectance (Table 1). The blue band is selected to maximise the blue-to-green ratio, R_{OCx} :

$$R_{OCx} = \log_{10} \left(\max \left(\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})} \right) \right). \tag{1}$$

The ratio of blue to green R_{rs} values is used as opposed to a single R_{rs} value in an attempt to normalize the data, as R_{rs} values in the blue part of the spectrum vary greatly with *chla*, whereas the green wavebands have a more limited range of variation. As *chla* increases, R_{rs} (443) will decrease and

Remote Sens. 2019, 11, 2609 6 of 29

become close to zero, reaching the noise level of the sensor. To avoid this issue, the algorithm switches to longer wavebands where the signal still correlates with *chla* and remains detectable. Additionally, longer wavelengths (e.g., 490 and 510) are less affected than the 443 nm band by variations in the concentration of coloured dissolved organic matter (i.e., yellow substances), and are therefore selected in the numerator of the ratio. Log10 *chla* is calculated using the following polynomial algorithm:

$$\log_{10}(chla) = a + b R_{OCx} + c R_{OCx}^2 + d R_{OCx}^3 + e R_{OCx}^4.$$
 (2)

NASA uses the method prescribed in O'Reilly et al. [10] to retrieve the optimal coefficients a, b, c, d, and e. This is an iterative method that constrains the slope and intercept of the linear regression between model and in situ chla to 1 and 0 respectively, minimizing the RMSLE (root mean squared logarithmic error, see Equation (12) in Section 2.5.1) and maximizing the coefficient of determination, r^2 . The comparison between band ratio models is then simplified since the slope, intercept, and consequently, the bias are all equal, and only the RMSLE and r^2 need to be evaluated. The in situ data used for the optimization of the original OCx algorithm is the global NASA Bio-Optical Marine Algorithm Dataset (NOMAD, version 2). Each sensor has a different set of wavelengths in the blue and green part of the spectrum (Table 1), such that the band ratio algorithm parameters are optimized for each sensor and are referred to as OC3M for MODIS, OC4 for SeaWiFS, and OC3V for VIIRS.

2.4.2. Regional Tuning of the OCx Algorithm

To remain consistent with the matchup exercise, the median value of the 3×3 pixel matrix of each waveband was computed to represent the $R_{rs}(\lambda)$ of the matchup. Note that the 443 nm waveband was removed from the potential blue wavebands in the band ratio as it was found to negatively affect the retrieval of *chla* (see discussion in Section 3.1). Several polynomial formulations from degrees one to four were tested and are hereafter referred to as "POLY1", "POLY2", "POLY3", and "POLY4" respectively. The optimal coefficients were retrieved using an iterative procedure to force the slope and intercept of the linear regression between satellite-derived and *in situ chla* to one and zero respectively as in O'Reilly et al. [10]. A 95% confidence interval for each optimal coefficient was computed using the *boot*() function in R with 2000 iterations [25,26]. For the sake of clarity and to investigate any potential benefits of adding higher-degree terms to the polynomial, the scores were computed for every polynomial degree, but only the results of the first and fourth-degree polynomials are discussed in the study. The parametrization and full results for the second and third polynomial are reported in Tables A1, A2 and A7 of Appendix A.

2.4.3. Original GSM

The GSM semi-analytical algorithm was developed by Garver and Siegel [27], and modified by Maritorena et al. [23]. The model uses a quadratic formulation (Equation (3)) to relate the underwater remote-sensing reflectance values to the ratio of total absorption, a (m^{-1}), and backscattering, b_b (m^{-1}) [28]. Note that the wavelength dependence in the following equations has been removed for the sake of legibility:

$$R_{rs,model}(0^-) = g_1\left(\frac{b_b}{a+b_b}\right) + g_2\left(\frac{b_b}{a+b_b}\right)^2,\tag{3}$$

where $g_1 = 0.0949$ and $g_2 = 0.0794$. Total absorption and backscattering are defined as follows:

$$a(\lambda) = a_w(\lambda) + a_{vh}(\lambda) + a_{dg}(\lambda),$$
 (4)

$$= a_w(\lambda) + chla \times a_{ph}^*(\lambda) + a_{dg}(443) \exp(-S(\lambda - 443)),$$
 (5)

Remote Sens. 2019, 11, 2609 7 of 29

and

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bv}(\lambda) \tag{6}$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda)$$

$$= b_{bw}(\lambda) + b_{bp}(443) \left(\frac{443}{\lambda}\right)^{Y},$$
(6)

where λ is the wavelength (nm). The term a corresponds to the total absorption term and is the sum of the absorption terms of pure seawater (a_w) , phytoplankton (a_{vh}) , and coloured detrital and dissolved organic materials (a_{dg}), which are combined in a single term given their similar spectral shapes; S expresses the exponential decrease of a_{dg} with increasing wavelength using the reference wavelength at 443 nm (i.e., a_{dg} (443)). The phytoplankton absorption term is defined as the concentration of *chla* multiplied by its specific absorption coefficient, a_{ph}^{\star} (i.e., absorption per unit *chla*), to account for its spectral variation. The total backscattering term is the sum of the backscattering terms of pure seawater (b_{bw}) and particulate matter (b_{bp}), the latter expressed as the product between a reference value (i.e., $b_{bp}(443)$) and the power-law-like decrease with increasing wavelength by the exponent Y. The coefficient a_{nh}^{\star} , given in Table A10 of Appendix A, was determined using in situ measurements of chla and phytoplankton absorption at the wavelengths of interest, and computing the mean value of absorption divided by chla across all available records within the AZMP database that spans from 1998 to 2014. The a_w coefficients were derived from tabulated values with a 2.5 nm resolution from 400 to 700 nm [29], and when necessary, interpolated to the corresponding wavelengths of each sensor. Similarly, the b_{bw} coefficients were derived from a 10 nm resolution table of values generated by Smith and Baker [30]. The spectral slope of yellow substances absorption S and the power-law decrease of particulate backscattering Y were optimized for use in the original algorithm to 0.02061 and 1.03373 respectively using a large simulated dataset [31]. The updated GSM models in this study tuned S and Y using the dataset collected from the NWA and NEP (see Section 2.4.4 for details).

 $R_{rs,satellite}(0^-)$ was retrieved by taking the median $R_{rs,satellite}(0^+)$ of the 3 × 3 matrix at each wavelength for a given match, then converting it from surface to underwater reflectance using Lee et al. [32]:

$$R_{rs,satellite}(0^{-}) = \frac{R_{rs}(0^{+})}{0.52 + 1.7R_{rs}(0^{+})}.$$
(8)

Non-linear least squares regression, using the Gauss–Newton algorithm of the nls() function [24] to minimize the difference between modelled and satellite-derived R_{rs} , was implemented with a maximum of 30 iterations to derive the remaining three unknowns in Equation (3): chla, the absorption coefficient of coloured detrital and dissolved material at the reference wavelength ($a_{dg}(443)$), and the backscattering coefficient of particulate matter at the reference wavelength ($b_{bp}(443)$). After retrieving these unknowns, a correction factor of 0.754188 was applied to the $a_{dg}(443)$ value according to Maritorena et al. [23]. The following constraints were then placed on each to ensure realistic values:

$$\begin{cases}
0.01 \le chla \le 64 \,\text{mg m}^{-3} \\
0.0001 \le a_{dg}(443) \le 2 \,\text{m}^{-1} \\
0.0001 \le b_{bp}(443)) \le 0.1 \,\text{m}^{-1}
\end{cases}$$
(9)

The original version of the model is abbreviated here as GSM_ORIG.

2.4.4. Regional Tuning of GSM

Similarly to the optimization of the OCx algorithm, $R_{rs}(\lambda)$ for a given satellite matchup was computed as the median of the corresponding 3×3 pixel matrix centered on the matchup pixel. Three modifications to the original GSM algorithm were tested. First, a new optimized exponent, P, was added to the chla term to correct for systematic underestimation at high chla and overestimation

Remote Sens. 2019, 11, 2609 8 of 29

at low *chla* (see Sections 3.2 and 3.3) that were potentially caused by the change of a_{ph}^{\star} with increasing *chla* (i.e., packaging effect [33]). The expression of the total absorption became:

$$a(\lambda) = a_w(\lambda) + chla^P \times a_{ph}^*(\lambda) + a_{dg}(443) \times \exp(-S(\lambda - 443)). \tag{10}$$

Second, the g parameters (see Equation (3) and Table A8) were optimized and spectral dependency was included as well as a new term g_3 (Equation (11)) using the synthetic dataset from the IOCCG working group on remote sensing of inherent optical properties [31]. Coefficients were retrieved at 10 nm intervals from 400 nm to 700 nm and interpolated to sensor-specific wavelengths when necessary.

$$R_{rs,model}(0^{-}) = g_1 \left(\frac{b_b}{a + b_b}\right) + g_2 \left(\frac{b_b}{a + b_b}\right)^{g_3}.$$
 (11)

Third, a sensitivity analysis was performed to compute the optimal exponents S (in the a_{dg} term), Y (in the b_{bp} term), and the new exponent P (in the a_{ph} term), which provided the best agreement between in situ and satellite-derived chla for the NWA and NEP regions. A total of 5355 (17 \times 35 \times 9) possible combinations of the three exponents were tested, ranging from 0.008 to 0.04 with an increasing step of 0.002 for S, 0.5 to 2.2 with an increasing step of 0.05 for Y, and 0.4 to 0.8 with an increasing step of 0.05 for P. These combinations of exponents were evaluated for all sensors and regions, as well as for both constant g values (see Equation (3)) and the new spectral g values (see Equation (11)). The best set of exponents was determined using a similar scoring method as that used for the algorithm performance evaluation (see Section 2.5.2), where each set of exponents was ranked according to a selected set of metrics. To remain consistent with the methods of optimizing band ratio coefficients, the slope and intercept, as well as RMSLE and r^2 , were selected for this test, and a score was assigned to each statistic based on its mean across all possible combinations and its value relative to other combinations. The optimal set of exponents was defined as the median of all sets with the highest sum of scores. At least 50% valid retrievals were required for each potential combination in the test, in order to avoid sets with low error values but with a poor predictive capacity. The optimal exponents are given in Table A9.

The modified versions of GSM were divided into three separate algorithms according to the changes that were made. All new versions included the exponent *P* on the *chla* term: 1) GSM_GC used the original *g* coefficients and regionally-tuned *P*, *S*, and *Y*, 2) GSM_GCGS used the spectral *g* coefficients in combination with the optimized *P*, *S*, and *Y* parameters from GSM_GC; and 3) GSM_GS used the spectral *g* coefficients optimized simultaneously with *P*, *S*, and *Y*. The results of GSM_GC and GSM_GS are discussed in this study to demonstrate improvement in *chla* retrieval when using the new spectrally-dependent *g* coefficients and the *P* exponent in the phytoplankton-dependent term, however only the scores of GSM_GCGS are included in the main text for simplicity's sake, and the full results in Appendix A (see Tables A1 and A2).

2.5. Performance Metrics

2.5.1. Statistical Models

Several sets of statistics were used for the evaluation of the algorithm. First, the total number of matchups (N) and the number of valid retrievals (n), were used to calculate the percentage of retrievals. Second, several formulations were used to express the error between $in\ situ$ and satellite-derived chla, including the mean error (μ_{error} , units of mg m⁻³), as well as the root mean squared logarithmic error (RMSLE):

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\log_{10}(C_i^*) - \log_{10}(C_i) \right)^2},$$
(12)

Remote Sens. 2019, 11, 2609 9 of 29

the mean log-transformed error (*MLE*):

$$MLE = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} \left(\log_{10}(C_i^*) - \log_{10}(C_i)\right)\right)},\tag{13}$$

and the mean magnitude of log-transformed error (MMLE),

$$MMLE = 10^{\left(\frac{1}{n}\sum_{i=1}^{n}\left|\log_{10}(C_{i}^{*}) - \log_{10}(C_{i})\right|\right)},$$
(14)

where C_i^* corresponds to satellite-derived *chla* of matchup i and C_i corresponds to *in situ chla* of matchup i. These three metrics are unitless as a result of the log transformation, and the latter two metrics have an ideal value of one, stemming from the reverse log transformation that converts them from linear to multiplicative space. Third, the intercept, slope, and coefficient of determination (r^2) were computed for the linear regression of log-transformed satellite-derived *chla* on log-transformed *in situ chla* using the Standard Major Axis method of Type 2 regression in the lmodel2() function [34]. Finally, a metric referred to as the "win ratio" based on Seegers et. al. [35], was also computed. The full dataset of $in \, situ/$ satellite matchups for a region/sensor was pared down to a new total value, T_{valid} , which only included matchups with valid satellite-derived *chla* for all algorithms. This facilitated comparisons between the errors of individual matchups for each algorithm. For a single $in \, situ/$ satellite pair, the algorithm that generated the lowest magnitude of error was the "winner". Each model's total number of "wins" in the set, $wins_{model}$, was then divided by T_{valid} :

$$win \, ratio_{model} = \frac{wins_{model}}{T_{valid}}.$$
 (15)

For example, if a dataset contains 500 matchups for which every algorithm generates a valid satellite *chla* concentration, and one of the algorithms (e.g., POLY1) produces the lowest error out of all algorithms for 70 of the 500 matchups, then the win ratio of POLY1 would be 70/500 = 0.14.

2.5.2. Scoring Method

The performance of all algorithms was evaluated using a modified version of the ranking approach developed by Brewin et al. [36], which relies on a slightly different set of metrics. Five statistics were selected as suggested in Seegers et al. [35] in an attempt to summarize the bias, accuracy, and precision of the model, as described below:

- 1. *MLE*, to account for the possible bias in the algorithm,
- 2. MMLE, to determine the accuracy of the algorithm,
- 3. r^2 of the linear regression, to test the precision or "goodness of fit" of the model,
- 4. percentage of valid retrievals, to ensure that a high score and low error are not reported as a result of a small number of retrievals, and
- 5. win ratio, to judge algorithm performance based on individual matches rather than a summary statistic.

The *RMSLE* was excluded from the scoring method in order to avoid redundant metrics that could skew the results by adding more emphasis to one element of the assessment (for example, bias) instead of giving all elements equal weight, however, its result is presented here for comparison with the literature. In our evaluation approach, all the metrics used were considered of equal weight. Each statistic was transformed to its magnitude of difference from the ideal value (for example, $r^2 = 0.6$ would be transformed to 0.4, as the ideal r^2 value is one) to give each the same reference point of zero so that lower statistics get higher scores. The mean of these transformed statistics was calculated across the results for all six algorithms. For a given algorithm, each of the five statistics was then scored as follows, taking into account their value relative to the mean and to the individual values of other algorithms:

Remote Sens. 2019, 11, 2609 10 of 29

• two points if the algorithm's transformed statistic was in the lowest 20% of the statistic's values across all algorithms (i.e., closer to the ideal value of zero) and below the mean,

- zero points if the algorithm's statistic was in the highest 20% of the statistic's values across all algorithms (i.e., further from the ideal) and above the mean, and
- one point otherwise

Statistical scores for each algorithm were then summed to give the algorithm's overall performance in comparison to others (Tables 3 and 4).

2.6. Temporal, Geographic and Phytoplankton Composition Influences on Algorithm Performance

The performance of the algorithms was also evaluated against environmental properties to detect any systematic bias or increase in uncertainties due to external factors. We studied the variability in the error (E_{chla}) as a function of time (defined as year + day of year/365) using boxplots generated by ggplot2() and $geom_boxplot()$ in R [37], given that the validation exercises span over several years for all three sensors and both regions. To examine the relationships between the uncertainties and other environment variables, a correlation analysis was performed on the magnitude of error ($|E_{chla}|$) between $in\ situ$ and satellite-derived chla against the variables listed below, using the shapiro.test() function to check that each subset followed a normal distribution, and the cor.test() function to retrieve the correlation coefficient and p-value [24]:

- 1. time, as defined above (units in decimal year),
- 2. day of year (units in days),
- 3. chlorophyll-a concentration (mg m^{-3}),
- 4. fucoxanthin to chlorophyll-a ratio (unitless),
- 5. latitude (decimal degrees),
- 6. bathymetry in coastal ocean (≤200 m depth, units in metres),
- 7. bathymetry in open ocean (>200 m depth, units in metres), and
- 8. distance (in metres) from central pixel to *in situ* samples, as we relax the constraint to 10 km.

The ratio of fucox to chla was selected to provide the ability to detect any bias due to change in phytoplankton composition, as the fucox pigment is an indicator of the presence of diatoms [38,39].

To test if the seasonal cycle induced different uncertainties we performed a Student's t-test on μ_{error} for each sensor and algorithm between spring and fall in the NWA, and similarly a One-Way ANOVA test between spring, summer and fall in the NEP. A separate t-test was also performed on μ_{error} between coastal and open ocean bathymetry subsets (as defined above). Again, the assumption of normality of each subset was checked using the shapiro.test() function and assumptions of equal variance between subsets were checked using the var.test() and fligner.test() functions for comparisons of two groups and three groups respectively. The tests for differences in means between subsets were conducted using the functions t.test() (for two groups) and aov() (three groups). Tukey's honest significant difference was performed on the ANOVA results from the NEP to determine which seasons generated mean errors that were statistically different from each other, using the TukeyHSD() function [24].

3. Results

3.1. Parameters of the Optimized Band Ratio Algorithms

Regionalization of the OCx and GSM algorithms was carried out using satellite-derived and in situ chla, which differs from the original algorithms where the coefficients of the models were derived using remote sensing reflectances and chla measured in situ. The regionally-tuned coefficients of the band ratio models remained consistent with the original parameters and notably, the first two coefficients (i.e., a and b, Table 2) were within the same order of magnitude and of the same sign as the original coefficients. In general, the first term a was higher in the regional versions of the algorithms. This can be explained by the overall underestimation of chla by the original OCx algorithms. The a and

Remote Sens. 2019, 11, 2609 11 of 29

b coefficients for the NEP for all sensors were closer to the original ones than for the NWA, which was consistent with the better performance of the original algorithms in the NEP than in the NWA (see Section 3.2 and 3.3). Regarding the coefficients for high degree polynomials (i.e., >= 2), there were no distinct patterns and the coefficients of the regional algorithms tend to be different and often opposite in sign from the original coefficients. The removal of the 443 nm band in the new band ratio algorithms may explain the differences between the original and regional coefficients of the algorithms. The numerator of the band ratio in the OCx algorithm varies between two or three possible wavebands depending on the sensor, where one of the options is the 443 nm band. Our analysis revealed that this band produced more outliers in the polynomial relationship between log-transformed in situ chla and the log-transformed band ratio than other numerator wavebands, which is particularly obvious in the NWA MODIS dataset (Figure 3). The blue boxes in Figure 3a,b show how the polynomial fit translates to a small range of modelled chla corresponding to a wide range of in situ chla. Removing the 443 nm waveband from the potential blue wavebands produced a tighter fit with the polynomial (Figure 3c,d). SeaWiFS and VIIRS NWA matchups datasets showed similar improvements after removing the 443 nm band, while the NEP datasets did not show significant changes after removing this band.

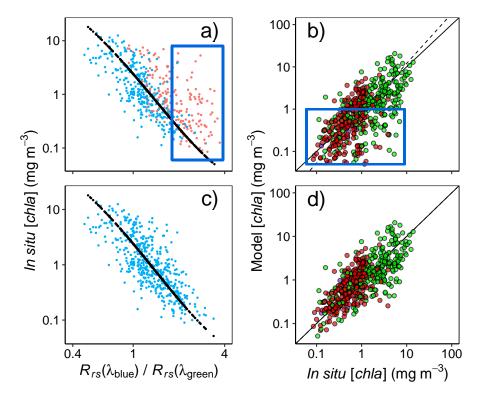


Figure 3. (a) *In situ chla* as a function of R_{rs} ratio for POLY4 including the 443 nm band, where the blue dots represent the band ratios that used the 488 nm band in the numerator and the pink dots used the 443 nm band in the numerator, (b) Satellite-derived versus *in situ chla* when including the 443 nm band in the ratio, where green and red circles represent spring and fall matchups respectively. (c) same as (a) but without including R_{rs} (443) in the POLY4 algorithm, (d) same as (b) but without including R_{rs} (443) in the POLY4 algorithm.

Remote Sens. 2019, 11, 2609 12 of 29

- 40 •		MODIS			SeaWiFS			VIIRS	
Coefficients	OC3M	POLY1	POLY4	OC4	POLY1	POLY4	OC3V	POLY1	POLY4
					NWA				
а	0.2424	0.36695	0.37925	0.3272	0.51664	0.51824	0.2228	0.43399	0.44786
b	-2.7423	-3.27757	-3.28487	-2.9940	-3.84589	-3.68431	-2.4683	-3.09652	-3.11091
С	1.8017	-	-0.75830	2.7218	-	-0.97401	1.5867	-	-0.77987
d	0.0015	-	1.49122	-1.2259	-	0.84875	-0.4275	-	1.42500
e	-1.2280	-	0.80020	-0.5683	-	0.77874	-0.7768	-	0.90445
					NEP				
а		0.24947	0.26575		0.41867	0.42516		0.31886	0.33055
b		-2.84152	-2.84142		-3.14708	-3.14271		-2.65010	-2.76455
С		-	-0.57938		-	-0.70269		-	-0.39595
d		-	0.74974		-	1.21802		-	1.52198
e		-	0.47743		-	1.59686		-	0.46509

Table 2. Coefficients of the OCx algorithms for each sensor and region.

3.2. Algorithms performance in the NWA

The statistic metrics remain consistent across the three sensors, with the regionally-tuned algorithms providing the best performance, as expected. In general, the polynomial-based approach performed better than the semi-analytical algorithms (Figure 4a and Table 3, see scores). It is noteworthy that the polynomial approaches yielded more matchups (96, 81 and 98% of possible matchups for MODIS, SeaWiFS and VIIRS respectively) than the semi-analytical methods (between 57 and 87% of possible matchups depending on sensors, global and regional methods). For the semi-analytical approach, the regional versions systematically provided more valid pixels that the original one.

Table 3. Statistics computed for OCx, POLY1, POLY4, Garver–Siegel Maritorena (GSM)_ORIG, GSM_GC, and GSM_GS algorithms for the Northwest Atlantic. Note that other versions of the algorithms OCx and GSM were tested but only included in Appendix A, so the "win ratio" column does not add up to one for each sensor as it does not include all values. Also recall that as a result of the log and reverse log transformations, the ideal value of *MLE* and *MMLE* is one. *N* is the total number of matchups as defined in Section 2.3, and *n* is the number of matchups that returned valid *chla* for the particular algorithm.

Algorithm	N	n	Intercept	Slope	r^2	μ_{error}	RMSLE	MLE	MMLE	Win Ratio	Score
]	MODIS					
OC3M	530	508	-0.067	0.839	0.45	-0.40	0.37	0.857	1.95	0.15	6
POLY1	530	508	-1.0×10^{-5}	1.00	0.57	0.081	0.34	1.00	1.85	0.037	6
POLY4	530	508	-4.2×10^{-6}	1.00	0.57	0.037	0.33	1.00	1.85	0.083	9
GSM_ORIG	530	354	-0.21	0.745	0.22	-0.84	0.47	0.644	2.16	0.32	2
GSM_GC	530	439	5.7×10^{-3}	0.995	0.56	0.13	0.34	1.01	1.88	0.10	5
GSM_GS	530	439	$7.3 imes 10^{-3}$	0.963	0.57	0.10	0.33	1.02	1.86	0.023	5
					9	eaWiFS					
OC4	416	336	-0.035	0.674	0.55	-0.58	0.32	0.928	1.75	0.25	7
POLY1	416	336	1.1×10^{-5}	1.00	0.62	-0.062	0.31	1.00	1.72	0.093	6
POLY4	416	336	-1.2×10^{-5}	1.00	0.62	-0.087	0.31	1.00	1.73	4.2×10^{-3}	5
GSM_ORIG	416	239	-0.18	0.734	0.28	-0.87	0.44	0.681	2.15	0.23	1
GSM_GC	416	282	0.025	0.957	0.64	0.057	0.29	1.06	1.69	0.11	5
GSM_GS	416	282	0.044	0.938	0.66	0.11	0.29	1.11	1.68	0.034	6
						VIIRS					
OC3V	176	172	-0.12	0.728	0.38	-0.45	0.37	0.831	1.90	0.13	3
POLY1	176	172	3.1×10^{-5}	1.00	0.55	-8.7×10^{-3}	0.33	1.00	1.82	0.035	6
POLY4	176	172	1.1×10^{-5}	1.00	0.55	-0.018	0.33	1.00	1.81	0.15	7
GSM_ORIG	176	144	-0.12	0.600	0.55	-0.45	0.31	0.906	1.78	0.24	5
GSM_GC	176	153	-0.036	0.959	0.56	-0.013	0.32	0.934	1.74	0.069	5
GSM_GS	176	153	0.018	1.04	0.58	0.054	0.32	1.02	1.75	0.21	6

Remote Sens. 2019, 11, 2609

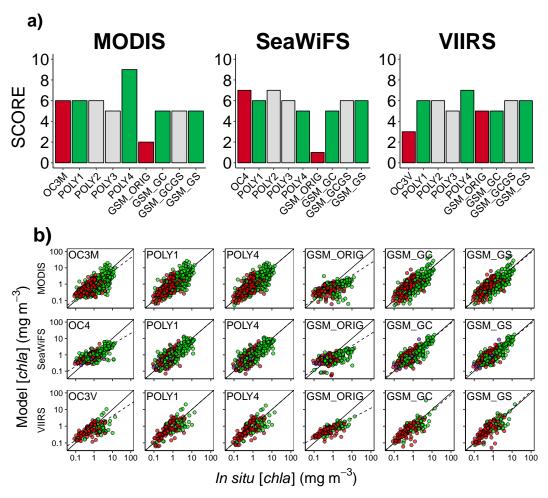


Figure 4. (a) Algorithms' scores for Moderate Resolution Imaging Spectrometer (MODIS), Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Visible Infrared Imaging Radiometer Suite (VIIRS) in the NWA based on the mean log-transformed error (MLE), mean magnitude of log-transformed error (MLE), and r^2 from the linear regression of satellite-derived on *in situ chla*, the percentage of retrievals, and the win ratio metric. Red bars correspond to the original algorithms, green bars correspond to regionally-tuned algorithms discussed in the text, grey bars corresponds to regionally-tuned algorithms presented only in Tables A1 and A2 of Appendix A. (b) Satellite-derived versus *in situ chla*, green solid circles correspond to the spring season and red solid circles correspond to the fall.

As shown in Table 3, all the generic algorithms (i.e., OCx and GSM_ORIG) had a slope of the linear regression of satellite-derived *chla* on *in situ chla* lower than 0.84 (0.839 for OC3M) and as low as 0.600 for VIIRS GSM_ORIG. Note that the slope of the linear regression of the tuned algorithms was forced to one. The regional algorithms also had a higher correlation coefficient than the original methods for all sensors, all algorithms. The *RMSLE* exhibited values (from 0.29 to 0.47) that were in agreement with values reported in the literature for regional algorithms for the retrieval of *chla* [40–42]. Tuned algorithms showed smaller *RMSLE* than original algorithms for both semi-analytical and band ratio approaches except for the GSM approaches for VIIRS. The *MLE* (Equation (13)) was lower than one for the original algorithms for all three sensors, as they have an overall negative bias, as seen in Figure 4b. After correcting the bias in the regional algorithms, the *MLE* was closer to one (particularly in the band ratio models) and the *MMLE* was reduced. In general, the improvement in the *MMLE* was greater for the GSM-type approaches than for the OCx-type approaches. The better performance of the regional algorithms is most noticeable in the mean error, which generally improved by a full order of magnitude. When the algorithms were applied to the identical subset of retrievals, GSM_ORIG showed the highest win ratio for MODIS and VIIRS, while narrowly losing to OC4 in the SeaWiFS

Remote Sens. 2019, 11, 2609 14 of 29

dataset. The regional linear and fourth-degree polynomial fits (POLY1 and POLY4) provided very similar results for all sensors, which is consistent with the study of Laliberté et al. [15] in the Gulf of Saint-Lawrence (Canada).

3.3. Algorithm Performance in the NEP

For the NEP, the original OCx algorithms exhibited a better linear fit than for the NWA (Table 4 and Figure 5b). For instance, OC3M showed a slope and intercept of the regression of satellite versus in situ data of 0.966 and 0.021 respectively against 0.839 and -0.067 for the NWA. The good performance of the original algorithms is also reflected in the correlation coefficients ($r^2 \ge 0.55$) for both original algorithms (OCx and GSM_ORIG) and all sensors, as they have higher values than for the NWA, with the exception of GSM_ORIG applied to VIIRS data where r^2 is 0.48 (Table 4). The regional algorithms, which aimed at reducing the RMSLE and forcing the slope and intercept to one and zero respectively, also improved the MLE. The RMSLE for all algorithms and sensors in the NEP are within the same order of magnitude as in the NWA except for the regional GSM models for SeaWiFS (i.e., GSM_GC and GSM_GS), which exhibited low RMSLE of 0.16 and 0.17 for GSM_GC and GSM_GS, respectively. As in the NWA, regional tuning generally improved the mean error, with the exception of the OCx and band ratio models in the MODIS dataset, where the magnitude of mean error increased, attesting to the good performance of the ocean colour model currently in use for this region and sensor. The semi-analytical algorithms (GSM-type) exhibited higher win ratios in general across all sensors, with GSM_ORIG showing the highest win ratio for MODIS and VIIRS (0.28 and 0.25 respectively), while GSM GS showed the highest win ratio for SeaWiFS. Ultimately, the simple linear band ratio model (POLY1) obtained the highest score across the three sensors but tied with POLY4 and the regional GSM models in the SeaWiFS dataset. It is worth noting that other variations of the band ratio models (POLY2 and POLY3) also performed well, with the third-degree polynomial (POLY3) outperforming POLY1 in the SeaWiFS set, and POLY2 obtaining a higher score than other polynomials in the NWA dataset (Table A1). Similarly, the MODIS NEP dataset contained an exception to the pattern of high-scoring GSM_GC and GSM_GS, where GSM_GCGS outperformed the two.

Table 4. Same as Table 3, for the NEP.

Algorithm	N	n	Intercept	Slope	r^2	μ_{error}	RMSLE	MLE	MMLE	Win Ratio	Score
]	MODIS					
OC3M	487	461	0.021	0.966	0.59	-0.016	0.37	1.05	1.87	0.12	5
POLY1	487	461	-5.3×10^{-6}	1.00	0.66	-0.21	0.34	1.00	1.78	0.11	7
POLY4	487	461	1.2×10^{-5}	1.00	0.67	-0.27	0.33	1.00	1.79	0.040	5
GSM_ORIG	487	279	-0.18	0.729	0.55	-0.96	0.37	0.708	1.81	0.28	3
GSM_GC	487	356	0.011	1.02	0.69	0.14	0.30	1.02	1.67	0.11	5
GSM_GS	487	355	2.1×10^{-3}	0.968	0.69	-0.013	0.29	1.01	1.65	0.040	5
					9	SeaWiFS					
OC4	45	40	-0.042	0.851	0.63	-0.34	0.30	0.897	1.65	0.15	3
POLY1	45	40	-1.3×10^{-5}	1.00	0.75	-3.8×10^{-4}	0.25	1.00	1.55	0.030	6
POLY4	45	40	5.3×10^{-6}	1.00	0.75	-5.5×10^{-3}	0.25	1.00	1.54	0.091	6
GSM_ORIG	45	33	-0.028	0.634	0.81	-0.42	0.23	0.959	1.56	0.15	4
GSM_GC	45	34	0.017	0.958	0.88	9.0×10^{-3}	0.16	1.04	1.30	0	6
GSM_GS	45	34	4.4×10^{-3}	0.972	0.87	-0.010	0.17	1.01	1.32	0.21	6
						VIIRS					
OC3V	342	332	-0.042	0.930	0.63	-0.36	0.34	0.902	1.78	0.17	6
POLY1	342	332	7.7×10^{-6}	1.00	0.69	-0.13	0.32	1.00	1.74	0.045	7
POLY4	342	332	1.4×10^{-5}	1.00	0.69	-0.17	0.31	1.00	1.74	0.063	6
GSM_ORIG	342	230	-0.10	0.725	0.48	-0.88	0.38	0.831	1.80	0.25	2
GSM_GC	342	265	0.021	1.01	0.68	0.095	0.29	1.05	1.60	0.13	6
GSM_GS	342	266	-3.4×10^{-3}	1.01	0.66	0.21	0.31	0.991	1.62	0.063	5

Remote Sens. 2019, 11, 2609 15 of 29

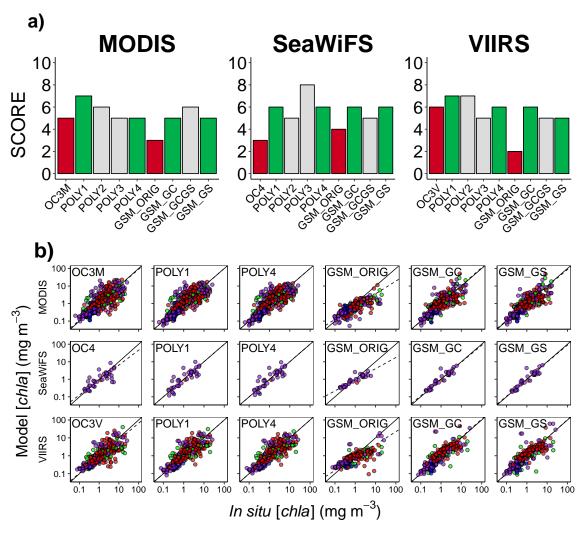


Figure 5. Same as Figure 4, for the NEP. Note the colour-coding for the solid circles in (**b**): blue corresponds to winter, green to spring, purple to summer, and red to fall.

3.4. Variation of Satellite-Derived Chla with Environmental Factors

3.4.1. Patterns in *Chla* Uncertainties with Phytoplankton Composition, Time, and Number of Retrievals

The correlation analysis described in Section 2.6 revealed that the magnitude of uncertainty in the retrieved chla ($|E_{chla}|$) was strongly positively correlated to chla concentration for all sensors, regions, and algorithms, with r^2 varying from 0.40 (VIIRS, GSM_GC, NWA) to 0.97 (GSM_ORIG, NWA, all sensors). GSM_GC presented an exception in the NEP SeaWiFS dataset with a weaker correlation between $|E_{chla}|$ and chla that can be explained by the very low number of matchups (Tables 5 and A2). Similarly, we found a relationship between $|E_{chla}|$ and the ratio of fucox to chla concentration for all algorithms, which was weaker in the regionally-tuned algorithms (Table 5) indicating that phytoplankton taxonomic composition has an impact on these types of algorithms, which can be reduced by the use of the regional models.

Remote Sens. 2019, 11, 2609 16 of 29

Table 5. Pearson's correlation coefficients (r) from the test for correlation between $|E_{chla}|$ and phytoplankton composition and time (see Section 2.6). * = statistically significant (p-value < 0.05).

]	MODIS	Se	aWiFS	V	/IIRS		MODIS	Sea	aWiFS		VIIRS
			N	WA					N	EP		
Algorithm	N	r	N	r	N	r	N	r	N	r	N	r
						Time						
OCx	508	-0.022	336	0.18 *	172	-0.10	461	0.17 *	40	0.20	332	-0.031
POLY1	508	1.1×10^{-3}	336	0.10	172	-0.087	461	0.17 *	40	0.14	332	6.7×10^{-4}
POLY4	508	3.4×10^{-3}	336	0.11 *	172	-0.088	461	0.17 *	40	0.15	332	0.020
GSM_ORIG	354	-0.058	239	0.27 *	144	-0.094	279	0.076	33	0.26	230	-0.083
GSM_GC	439	-0.14	282	0.16 *	153	-0.043	356	6.5×10^{-4}	34	0.25	265	-0.033
GSM_GS	439	-0.14	282	0.15 *	153	-0.012	355	-8.1×10^{-3}	34	0.21	266	-0.10
					D	ay of year						
OCx	508	-0.41 *	336	-0.35 *	172	-0.46 *	461	0.14 *	40	0.15	332	0.14 *
POLY1	508	-0.33 *	336	-0.26*	172	-0.39*	461	0.15 *	40	0.11	332	0.15 *
POLY4	508	-0.35*	336	-0.26*	172	-0.41*	461	0.15 *	40	0.11	332	0.16 *
GSM_ORIG	354	-0.48*	239	-0.40*	144	-0.35*	279	0.065	33	0.19	230	0.13 *
GSM_GC	439	-0.22*	282	-0.21*	153	-0.25*	356	0.12 *	34	0.039	265	0.088
GSM_GS	439	-0.22 *	282	-0.22*	153	-0.39 *	355	0.12 *	34	0.087	266	0.095
						[Chla]						
OCx	508	0.78 *	336	0.91 *	172	0.96 *	461	0.71 *	40	0.68 *	332	0.82 *
POLY1	508	0.56 *	336	0.70 *	172	0.70 *	461	0.79 *	40	0.46 *	332	0.81 *
POLY4	508	0.61 *	336	0.72 *	172	0.73 *	461	0.80 *	40	0.46 *	332	0.82 *
GSM_ORIG	354	0.97 *	239	0.97 *	144	0.97 *	279	0.95 *	33	0.87 *	230	0.91 *
GSM_GC	439	0.52 *	282	0.64 *	153	0.40 *	356	0.59 *	34	0.28 *	265	0.57 *
GSM_GS	439	0.52 *	282	0.66 *	153	0.73 *	355	0.64 *	34	0.42 *	266	0.63 *
					[Fu	cox]/[chla]					
OCx	508	0.37 *	336	0.36 *	172	0.53 *				<u> </u>		
POLY1	508	0.27 *	336	0.34 *	172	0.43 *						
POLY4	508	0.29 *	336	0.34 *	172	0.45 *						
GSM_ORIG	354	0.48 *	239	0.38 *	144	0.54 *						
GSM_GC	439	0.26 *	282	0.34 *	153	0.27 *						
GSM_GS	439	0.27 *	282	0.35 *	153	0.38 *						

There were no significant long-term trends of error magnitude in the time series for all sensors and both regions (Table 5). However, there was a negative correlation between the magnitude of error and the day of year across all sensors in the NWA, which is also apparent in the larger spread of error (the interquartile range) occurring during the first half of each year (1.5 mg m⁻³ on average across sensors for both POLY4 and GSM_GS) than during the last half (0.59 and 0.63 for POLY4 and GSM_GS respectively), with a general negative bias in the spring across the whole time series (Figure 6 and Table A6). The differences in the spread of error between the first and last halves of the year is generally less noticeable in the NEP (Figure 7), where there is a smaller positive correlation between error magnitude and day of year, corresponding to a larger average spread of error in the second half of each year (1.0 and 0.57 mg m⁻³ for POLY4 and GSM_GS respectively) than in the first half (0.51 and 0.34). This can possibly be attributed in part to the timing of the samples, which were spread out over the year instead of concentrated in the spring and fall as in the NWA.

Remote Sens. 2019, 11, 2609 17 of 29

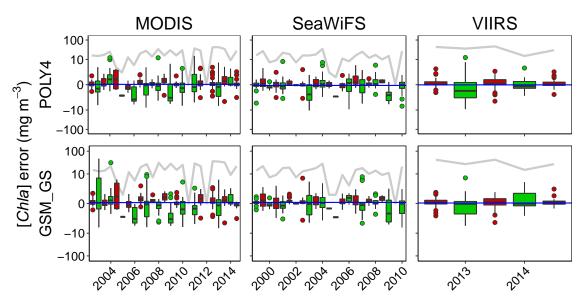


Figure 6. Model residuals across the full-time series for the POLY4 and GSM_GS algorithms in the NWA for MODIS, SeaWiFS and VIIRS. Blue lines represent μ_{error} , and the grey lines indicate the number of valid retrievals. Green boxes represent the first half of the year (January to June) and red boxes represent the second half (July to December). Note the axis is asinh-transformed.

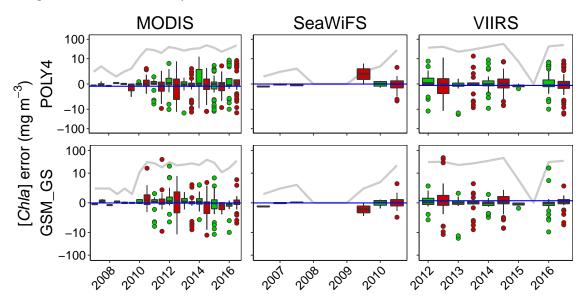


Figure 7. Same as Figure 6, for the NEP.

A t-test and ANOVA also revealed that in the NWA, the average μ_{error} across algorithms in the spring (-0.51, -0.60, and -0.60 mg m $^{-3}$ for MODIS, SeaWiFS, and VIIRS respectively) were generally significantly different (opposite in sign and higher in magnitude) than in the fall (0.20, 0.19, and 0.032) (see Table A3), but there were no significant differences in the mean error between seasons in the NEP (Table A4). We did not find any correlation between the number of matchups and the magnitude of the uncertainties for a given season/sensor/region except when only a very low number of matchups was available (e.g., MODIS in spring 2005 in the NWA, 2 matchups, Figure 6) and one bad matchup could negatively weight the entire matchup dataset. In the NEP (Figure 7), there was a large number of outliers spread evenly across spring and fall seasons, unlike the NWA, where spring blooms generated the highest spread of error. Only a few data were available for SeaWiFS in the NEP, which were concentrated mainly in 2009 and 2010, and therefore the correlation and ANOVA results for this region/sensor dataset were less reliable. All variations of GSM also returned fewer valid retrievals

Remote Sens. 2019, 11, 2609 18 of 29

than the band ratio algorithms, mainly due to negative remote sensing reflectances at either end of the spectrum, which prevented the model from generating a valid fit (Table 6).

Table 6. Missing retrievals by GSM_GS algorithm for each region/sensor combination, subdivided by the cause of lost data, with the percentage of lost retrievals given for each cause. $R_{rs}(41X)$ is the remote sensing reflectance value at 412 nm for MODIS and SeaWiFS, and 410 nm for VIIRS. $R_{rs}(6XX)$ varies from 645 to 671 nm depending on sensor. Note that there were no instances of negative model *chla*, unlike $a_{dg}(443)$ and $b_{bp}(443)$. "Unrealistic unknowns" occur when *chla*, $a_{dg}(443)$, or $b_{bp}(443)$ are outside their acceptable ranges, given in Equation (9). Multiple issues were a combination of $a_{dg}(443) < 0$ and *chla* above the acceptable range, with two instances where $b_{bp}(443)$ was also too high (>0.18).

	MO	DDIS	Sea	WiFS	V]	IIRS	MO	DDIS	Se	aWiFS	V]	IIRS
Cause of Missing Pixel			N	WA					N	EP		
	#	%	#	%	#	%	#	%	#	%	#	%
$R_{rs}(41X) < 0$	34	49.3	28	51.9	7	36.8	66	62.3	0	0	26	39.4
$R_{rs}(6XX) < 0$	5	7.25	3	5.56	1	5.26	8	7.55	2	33.3	3	4.55
$a_{dg}(443) < 0$	27	39.1	20	37.0	6	31.6	17	16.0	4	66.7	18	27.3
$b_{bv}(443) < 0$	0	0	0	0	0	0	1	0.940	0	0	0	0
Unrealistic unknowns	3	4.35	2	3.70	2	10.5	7	6.60	0	0	3	4.55
Multiple issues	0	0	0	0	0	0	7	6.60	0	0	2	3.03
Unexplained	0	0	1	1.85	3	15.8	0	0	0	0	14	21.2
Total	69		54		19		106		6		66	

3.4.2. Uncertainties Related to Geographic Considerations

The correlation analysis of $|E_{chla}|$ on latitude, bathymetry and distance from *in situ* to satellite measurements was carried out to investigate if these variables could explain some of the biases and discrepancies observed between the *in situ* and satellite-derived *chla* (Table 7). There were no correlations observed between $|E_{chla}|$ and latitude, which is an interesting result given the range of latitudes in the NWA dataset from approximately 38° to 61°N. Possible exceptions exist in the NWA VIIRS and NEP SeaWiFS datasets, but as mentioned in the previous section, the SeaWiFS dataset had very few data, and the high latitude data in the VIIRS dataset were all retrieved during the spring bloom (see Figure 1a,b), when *chla* concentration would typically be higher and thus $|E_{chla}|$ would likely also be higher as the two were collinear (see Section 3.4.1).

There were low negative correlations between $|E_{chla}|$ and open ocean bathymetry in the NWA MODIS dataset, and larger negative correlations in the NEP, which again can be explained by the collinearity of *chla* concentration and $|E_{chla}|$ and the fact that in the open ocean subset, bathymetry was also negatively correlated with *chla* concentration (r=-0.47). As for seasonal bias, the *t*-test on the coastal and offshore datasets in the NWA revealed that the mean uncertainties for both datasets were statistically different in the regional band ratio models for MODIS and SeaWiFS, with μ_{error} ranging from 0.14 to 0.42 mg m⁻³ for the coastal regions and -0.31 to -0.20 in open ocean, indicating that this family of algorithms systematically underestimated *chla* offshore and overestimates in shallower water, while the difference was less significant with the GSM-type algorithms (see Table A5). The pattern reversed in the NEP, where the regional band ratio underestimated coastal *chla* (-0.90 to -0.77 mg m⁻³) and overestimated *chla* in the open ocean (0.076 to 0.27), again ignoring the SeaWiFS dataset due to insufficient data.

Distance between a satellite matchup and the *in situ* measurement was up to 10 km to maximize the number of matchups, though over half the matchups in each region/sensor subset were within 2 km of the measurement. Regression of $|E_{chla}|$ on the distance did not reveal a relationship (|r| < 0.1) for any datasets except NWA VIIRS GSM_ORIG, and the small NEP SeaWiFS dataset, suggesting that in open waters, where spatial variations in *chla* are small (mesoscale) compared to the coastal

Remote Sens. 2019, 11, 2609

environment, the stringent criteria of pixel to *in situ* collocation can be relaxed to increase the number of matchups.

Table 7. Pearson's correlation coefficients (r) for the test for correlation between $|E_{chla}|$ and geographic environmental factors (see Section 2.6). * = statistically significant (p-value < 0.05).

		MODIS		SeaWiFS		VIIRS		MODIS	Se	eaWiFS	V	'IIRS
				NWA					N	EP		
Algorithm	N	r	N	r	N	r	N	r	N	r	N	r
					La	titude						
OCx	508	0.095 *	336	-0.077	172	0.37 *	461	0.011	40	0.23	332	0.022
POLY1	508	0.076	336	-0.022	172	0.33 *	461	-4.3×10^{-3}	40	0.19	332	0.019
POLY4	508	0.079	336	-0.031	172	0.34 *	461	-0.010	40	0.20	332	0.016
GSM_ORIG	354	0.049	239	-0.12	144	0.30 *	279	-0.076	33	0.13	230	-0.075
GSM_GC	439	-0.017	282	-0.014	153	0.21 *	356	3.2×10^{-3}	34	0.45 *	265	-0.076
GSM_GS	439	-0.018	282	-0.021	153	0.35 *	355	4.2×10^{-3}	34	0.38 *	266	-0.084
					Ope	n ocean						
OCx	276	-0.16 *	171	-0.13	98	-0.029	278	-0.45 *	30	-0.41 *	197	-0.41 *
POLY1	276	-0.18*	171	-0.13	98	-0.059	278	-0.44*	30	-0.38*	197	-0.45*
POLY4	276	-0.18*	171	-0.13	98	-0.059	278	-0.45*	30	-0.39*	197	-0.46*
GSM_ORIG	212	-0.17*	132	-0.14	87	-0.033	202	-0.32*	25	-0.53*	161	-0.28*
GSM_GC	244	-0.15*	148	-0.10	89	3.9×10^{-3}	238	-0.33*	26	-0.54*	177	-0.32*
GSM_GS	244	-0.15 *	148	-0.10	89	0.074	237	-0.35 *	26	-0.47*	177	-0.33*
					Coast	tal ocean						
OCx	232	-9.3×10^{-3}	165	-0.051	74	0.13	186	-0.083	10	-0.59	137	-0.090
POLY1	232	-0.011	165	-0.11	74	0.12	186	-0.10	10	-0.37	137	-0.090
POLY4	232	-0.024	165	-0.10	74	0.11	186	-0.11	10	-0.36	137	-0.079
GSM_ORIG	142	0.032	107	-1.5×10^{-3}	57	0.18	78	-0.091	8	-0.05	69	-0.051
GSM_GC	195	-0.019	134	0.052	64	0.22	119	-1.9×10^{-3}	8	-0.48	88	0.037
GSM_GS	195	-0.017	134	0.054	64	0.18	119	-8.3×10^{-3}	8	-0.36	89	0.015
				Distance	to in	situ measurei	nent					
OCx	508	-0.033	336	0.037	172	-0.038	461	0.036	40	-0.11	332	0.020
POLY1	508	-0.054	336	0.020	172	-0.085	461	0.046	40	-0.14	332	0.037
POLY4	508	-0.049	336	0.026	172	-0.085	461	0.050	40	-0.13	332	0.052
GSM_ORIG	354	-0.057	239	-0.15 *	144	-0.026	279	-6.3×10^{-3}	33	-0.32	230	-0.010
GSM_GC	439	-0.055	282	-0.077	153	-0.045	356	0.047	34	-0.32	265	0.076
GSM_GS	439	-0.059	282	-0.075	153	-0.072	355	0.057	34	-0.35*	266	0.012

4. Discussion

Satellite-derived chla concentration remains the most used product to infer global scale information on the status of the marine ecosystem, and non-specialists represent an important fraction of ocean colour data users. Here we have assessed the performance of two algorithms currently implemented in NASA's SeaDAS software (https://seadas.gsfc.nasa.gov) in Canadian waters (Atlantic and Pacific Oceans) to inform on biases associated with these algorithms in those regions. In addition, we carried out modifications of these original models by optimizing their parameterization and formulation to correct for regional bias in the NWA and NEP. Both OCx and GSM exhibited a negative overall bias particularly in the NWA, where low chla were overestimated and high concentrations were underestimated, which resulted in a linear fit of satellite-derived on in situ chla with a slope lower than one. This bias was partially corrected in the modified versions of the models, which also improved the tightness of the relationship as indicated by higher r^2 than for the original algorithms, but at the expense of the number of individual retrievals with the lowest magnitude of error between different algorithms ("win ratio", Tables 3 and 4). Significantly different means in chla concentration between spring and fall in the NWA (see Section 3.4.1) suggest that the dataset could be further subdivided by season, but at the expense of a synoptical approach. It would also require a larger matchup dataset to achieve statistical significance.

Our approach differed from NASA's in that the algorithms were fitted directly between satellite-derived and *in situ chla*, under the assumption that the satellite R_{rs} were reliable, while

Remote Sens. 2019, 11, 2609 20 of 29

the original algorithms were developed using both in situ radiometric and chla measurements. Furthermore, the regional parameters were optimized, tested, and compared to the original algorithms using the entire dataset for a given region and sensor, rather than extracting a subset for training and using the remainder for testing. Confidence intervals for the polynomial algorithm coefficients were retrieved using a bootstrap method that subsamples the original dataset to compute 2000 different sets of coefficients (see Section 2.4.2), revealing interval widths ranging from 0.08 to 1.26 for the first three coefficients and 0.38 to 2.39 for the last two, which had the smallest impact on the shape of the polynomial in the region defined by the band ratio values. This suggests that the coefficients derived from the full dataset are representative of any subsets. For the GSM algorithms, the sensitivity study on combinations of spectral variations in P, S, and Y (see Section 2.4.4) was performed for each region, sensor, and set of g coefficients (spectral or constant), and the median of the subset of highest-scoring combinations was selected. For this reason, we can assume that using the entire dataset or a subset for each combination would have led to similar results after retrieving the median of the optimal combinations. Finally, the small size of some of the datasets, particularly the NEP SeaWiFS set, presented a challenge in developing a reliable set of parameters spanning a wide range of values for in situ chla and the environmental variables that were tested for correlations with algorithm error. For these reasons, the regional tuning of the algorithms was carried out on the full datasets rather than subsets.

One of the main results of our studies was the identification of the 443 nm waveband as a strong contributor to the overall uncertainty in the OCx algorithms. Removing this waveband from the OCx algorithms and iteratively forcing the slope and intercept of the linear model between satellite-derived and *in situ chla* to one and zero respectively provided better results than the original algorithms (Table 3). Increasing the degree of the polynomial improved the results for some combinations of region and sensor (see scores in Tables 3 and 4), but overall offered minimal improvements. The negative bias of the original GSM model was corrected by introducing a new exponent, P, on the chlorophyll term, which lowered satellite-derived *chla* concentrations <1 mg m⁻³ and increased concentrations >1 mg m⁻³. This exponent P was derived in combination with the exponents on the a_{dg} and b_{bp} terms, giving the best set of spectral slopes for each region and sensor. Inclusion of spectral dependence in the g coefficients in Equation (11) slightly improved the results, however, optimization of the spectral slopes of the phytoplankton and yellow substances absorption terms, as well as the particulate backscattering term, had a more positive impact. This finding highlights the importance of accounting for regional properties of absorption and backscattering terms, and particularly their spectral dependence. Overall, the GSM-type algorithms provided the highest win ratio (i.e., the highest number of matchups with the smallest chla uncertainties), however, this good performance was hampered by the lower number of retrievals compared to the polynomial formulations (Tables 3 and 4). This limitation has to be considered according to the end-user's applications of the chla product.

Uncertainties in Satellite-derived products are currently an active field of research (IOCCG report on Uncertainties in Ocean Colour Remote Sensing, in preparation) that remains highly complex given the possible sources of uncertainties associated with satellite observations that include, among others, uncertainties in radiometry (and calibration), atmospheric corrections, data binning, and bio-optical algorithms. For instance, it has been shown that at the global scale, the OCx algorithm contains an inherent uncertainty of about 35% [43]. The difference in scales between satellite and *in situ* measurements also creates another source of uncertainty, as *in situ* data refer to about two litres or less of seawater at a discrete depth while satellites integrate a volume of tens of millions of litres of seawater. Note that uncertainties in HPLC-derived *chla* were partially accounted for in the linear model comparing the results by using Type 2 linear regression as described in Section 2.5.1, which minimizes the sum of areas of triangles between each point and the regression line, assuming that both variables contain errors. Here, we have addressed discrepancies between satellite-derived *chla* and *in situ chla* and included temporal and spatial effects. For the two areas of interest, namely the Northwest Atlantic and the Northeast Pacific, we did not find significant patterns of temporal drift,

Remote Sens. 2019, 11, 2609 21 of 29

thanks to regular reprocessing of ocean colour products by NASA following a vicarious calibration exercise [44] that ensures stable measurements of the radiative field with time. We did not find spatial patterns, with the exception of minor positive correlations between latitude and magnitude of error in the NWA VIIRS dataset, and negative correlations between open ocean bathymetry and error magnitude in the NEP datasets, which are likely due to changes in chla concentration that affect the degree of error. Seasonal biases were observed with an overall underestimation of chla in the spring and a small overestimation in the fall. This could be explained by the change in phytoplankton community composition (Table 5), such that when the ratio of fucoxanthin, a marker pigment for diatoms, to chla concentration increased, the magnitude of error in chla retrieval also increased. This pattern was observed across all algorithms, but weaker in the regionally-tuned algorithms. GSM-type algorithms offered the greatest improvements, in particular, GSM_GC in the VIIRS dataset where the correlation coefficient of $|E_{chla}|$ and the ratio of fucox to *chla* was reduced from 0.54 to 0.27. This demonstrates the effectiveness of regional algorithms at addressing changes in phytoplankton absorption properties, and notably the modified GSM algorithms, which include an exponent on the phytoplankton absorption term to account for changes in absorption properties with changes in phytoplankton community composition in the Northwest Atlantic [45–47]. However, the fact that the fucox to chla ratio impacted the regional algorithms for SeaWiFS (i.e., both OCx-like and GSM-like) to a similar degree as the original models suggests that the 510 nm band, which is used across all models for the SeaWiFS sensor, can contain information on the presence of diatoms, as highlighted in Sathyendranath et al. [48] who developed an algorithm to discriminate diatoms from other phytoplankton types for SeaWiFS.

Finally, another notable source of uncertainty is the atmospheric correction method applied to NASA's level 2 images, based on the "black pixel assumption", which assumes that scattering in the near-infrared bands from approximately 670 nm to 865 nm [49] is negligible, thus the ocean appears black. This was a reasonable assumption for case 1 waters [50], but optically-complex case 2 waters can contain elements that contribute to the scattering in those bands, giving an inaccurate correction of remote sensing reflectance values used in *chla* algorithms [51]. Given that many matchups in this report were near coastal or shelf areas, there could have been a degree of error associated with poor atmospheric correction. As discussed in Section 3.4.1, many satellite matchups did contain negative R_{rs} in a waveband near one end of the visible spectrum. These matchups were incapable of computing valid *chla* concentrations for the GSM algorithms but were included in the band ratio algorithms.

5. Conclusions

This comparison exercise between satellite-derived and *in situ chla* concentration measurements for three ocean colour sensors (i.e., SeaWiFS, MODIS and VIIRS), two widely-used algorithms based on different approaches (i.e., band ratio and semi-analytical) and for two regions (i.e., Northwest Atlantic and Northeast Pacific) revealed systematic bias (slopes varying from 0.60 to 0.84 in the NWA and from 0.63 to 0.97 in the NEP) and relatively large uncertainties. While regional tuning permitted correction for the biases, the uncertainties remained substantial, attesting to the fact that natural variations of phytoplankton optical properties, the presence of other optically active components, and accuracy of the atmospheric correction will challenge accurate retrieval of *chla* concentration. In the Northwest Atlantic, the removal of the 443 nm band in the regional band ratio algorithm improved the performance of the OCx-type algorithms. It is noteworthy that while band ratio algorithms provided the most matchups between satellite and *in situ* data, the GSM-type approach appeared to be most accurate thanks to its ability to decipher the phytoplankton signal from yellow substances, but its performance was hindered by the significantly lower number of retrievals, mainly due to negative radiances present in the blue or red part of the visible spectrum, perhaps due to poor atmospheric corrections.

Remote Sens. 2019, 11, 2609 22 of 29

Author Contributions: E.D. conceptualized and led the project, E.D. and S.C. developed the methodology. S.C. wrote most of the code and carried out all the statistical analyses. B.D. contributed to the coding and provided technical support. A.P. provided the dataset for the Northeast Pacific. S.C. and E.D. drafted the manuscript, which was edited by B.D. and A.P.

Funding: This research was funded by MEOPAR under the Observation Core Project and the Strategic Program for Ecosystem-Based Research and Advice (SPERA) from Fisheries and Oceans Canada.

Acknowledgments: The authors would like to thank NASA's Ocean Biology Processing Group (OBPG) for supplying a large up-to-date satellite ocean colour dataset available for public use (https://oceancolor.gsfc.nasa.gov/), as well as detailed information on the current ocean colour algorithms.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NWA Northwest Atlantic NEP Northeast Pacific SeaWiFS Sea-viewing Wide Field-of-view Sensor **MODIS** Moderate Resolution Imaging Spectrometer **VIIRS** Visible Infrared Imaging Radiometer Suite **RMSLE** Root mean squared logarithmic error MLEMean log-transformed error **MMLE** Mean magnitude of log-transformed error OCx Ocean Colour X Garver-Siegel Maritorena **GSM** N total number of matchups number of valid retrievals mean error μ_{error} coefficient of determination E_{chla} error in chlorophyll-a concentration retrieval $|E_{chla}|$ magnitude of the error in chlorophyll-a concentration retrieval

Appendix A. Results Omitted from the Main Text

Table A1. Statistics of extra algorithms not presented in the text, as in Tables 3 and 4.

Algorithm	N	n	Intercept	Slope	r ²	μ_{error}	RMSLE	MLE	MMLE	Win Ratio	Score
					NWA	-MODIS					
POLY2	530	508	-1.2×10^{-5}	1.00	0.57	1.3×10^{-3}	0.33	1.00	1.85	0.10	6
POLY3	530	508	1.4×10^{-5}	1.00	0.57	0.036	0.33	1.00	1.85	0.011	5
GSM_GCGS	530	439	0.037	0.959	0.57	0.23	0.33	1.09	1.86	0.17	5
					NWA-	—SeaWiFS					
POLY2	416	336	-5.5×10^{-7}	1.00	0.62	-0.10	0.31	1.00	1.74	0.068	7
POLY3	416	336	$1.9 imes 10^{-5}$	1.00	0.62	-0.077	0.31	1.00	1.73	0.076	6
GSM_GCGS	416	283	0.065	0.931	0.66	0.19	0.29	1.17	1.69	0.14	6
					NWA	A—VIIRS					
POLY2	176	172	5.8×10^{-5}	1.00	0.55	-0.11	0.33	1.00	1.83	0.090	6
POLY3	176	172	3.9×10^{-5}	1.00	0.55	-0.029	0.33	1.00	1.81	6.9×10^{-3}	5
GSM_GCGS	176	154	3.7×10^{-3}	0.919	0.57	0.15	0.30	1.04	1.72	0.069	6

Remote Sens. 2019, 11, 2609 23 of 29

Table A1. Cont.

Algorithm	N	n	Intercept	Slope	r ²	μ_{error}	RMSLE	MLE	MMLE	Win Ratio	Score
					NEP-	MODIS					
POLY2	487	461	-1.9×10^{-5}	1.00	0.67	-0.36	0.33	1.00	1.80	0.098	6
POLY3	487	461	1.4×10^{-5}	1.00	0.67	-0.33	0.33	1.00	1.79	0.040	5
GSM_GCGS	487	355	0.039	0.975	0.69	0.16	0.29	1.10	1.66	0.16	6
					NEP—	SeaWiFS					
POLY2	45	40	2.9×10^{-5}	1.00	0.75	-0.036	0.25	1.00	1.55	0	5
POLY3	45	40	2.2×10^{-7}	1.00	0.75	-0.071	0.25	1.00	1.55	0.21	8
GSM_GCGS	45	34	0.038	0.941	0.88	0.083	0.16	1.09	1.32	0.15	5
					NEP-	-VIIRS					
POLY2	342	332	-4.9×10^{-5}	1.00	0.69	-0.23	0.31	1.00	1.74	0.080	7
POLY3	342	332	1.7×10^{-5}	1.00	0.69	-0.18	0.31	1.00	1.73	0.022	5
GSM_GCGS	342	270	0.026	0.997	0.56	0.028	0.35	1.06	1.69	0.17	5

Table A2. Pearson's correlation coefficients (r) from the test for correlation between $|E_{chla}|$ and environmental factors (see Section 2.6) for model versions not presented in the text. * = statistically significant (p-value < 0.05).

	1	MODIS	Se	aWiFS	V	'IIRS]	MODIS	Se	eaWiFS	V	'IIRS
			N	WA					N	NEP		
Algorithm	N	r	N	r	N	r	N	r	N	r	N	r
					La	titude						
POLY2	508	0.081	336	-0.035	172	0.38 *	461	-0.014	40	0.21	332	0.010
POLY3	508	0.079	336	-0.028	172	0.34 *	461	-0.011	40	0.22	332	0.016
GSM_GCGS	439	-0.018	283	-0.020	155	0.19 *	355	2.9×10^{-3}	34	0.40 *	270	-0.092
					Оре	n ocean						
POLY2	276	-0.18*	171	-0.13	98	-0.068	278	-0.44*	30	-0.41*	197	-0.46*
POLY3	276	-0.18*	171	-0.13	98	-0.059	278	-0.44 *	30	-0.44*	197	-0.46*
GSM_GCGS	244	-0.14 *	149	-0.13	89	-0.016	237	-0.36 *	26	-0.45 *	179	-0.35*
					Coas	tal ocean						
POLY2	232	-0.024	165	-0.097	74	0.067	186	-0.090	10	-0.39	137	-0.076
POLY3	232	-0.022	165	-0.10	74	0.11	186	-0.091	10	-0.41	137	-0.078
GSM_GCGS	195	-0.025	134	0.055	66	0.20	119	8.3×10^{-3}	8	-0.48	91	0.013
				Distanc	e to in	situ meas	ureme	nt				
POLY2	508	-0.047	336	0.027	172	-0.080	461	0.048	40	-0.14	332	0.053
POLY3	508	-0.049	336	0.024	172	-0.085	461	0.049	40	-0.13	332	0.053
GSM_GCGS	439	-0.060	283	-0.077	155	-0.050	355	0.056	34	-0.35 *	270	0.057
					-	Гіте						
POLY2	508	1.9×10^{-3}	336	0.11 *	172	-0.094	461	0.17 *	40	0.17	332	0.015
POLY3	508	2.8×10^{-3}	336	0.11	172	-0.089	461	0.17 *	40	0.21	332	0.020
GSM_GCGS	439	-0.14 *	283	0.14 *	155	-0.040	355	-0.020	34	0.22	270	-0.031
					Day	of year						
POLY2	508	-0.35*	336	-0.26*	172	-0.42*	461	0.15 *	40	0.12	332	0.16 *
POLY3	508	-0.35 *	336	-0.25*	172	-0.41*	461	0.15 *	40	0.12	332	0.16 *
GSM_GCGS	439	-0.20 *	283	-0.21*	155	-0.20*	355	0.11 *	34	0.088	270	0.094
					[-	Chla]						
POLY2	508	0.64 *	336	0.74 *	172	0.87 *	461	0.84 *	40	0.49 *	332	0.84 *
POLY3	508	0.61 *	336	0.72 *	172	0.75 *	461	0.83 *	40	0.54 *	332	0.82 *
GSM_GCGS	439	0.49 *	283	0.65 *	155	0.33 *	355	0.58 *	34	0.39 *	270	0.59 *
					[Гисо	x] / [chla]						
POLY2	508	0.30 *	336	0.34 *	172	0.47 *						
POLY3	508	0.29 *	336	0.34 *	172	0.45 *						
GSM_GCGS	439	0.25 *	283	0.34 *	155	0.22 *						

Remote Sens. 2019, 11, 2609 24 of 29

Table A3. Results of the *t*-test performed on μ_{error} of seasonal subsets in the NWA to test for statistically significant differences in mean error between spring and fall. N = size of subset. * = statistically significant (p-value < 0.05).

Algorithm	$N_{ m spring}$	N_{fall}	$\mu_{error}^{ m spring}$	$\mu_{error}^{\mathrm{fall}}$	<i>p</i> -Value
			MOI	DIS	
OCx	254	239	-0.86	0.074	3.3×10^{-7} *
POLY1	254	239	-0.077	0.26	0.14
POLY4	254	239	-0.18	0.27	0.024 *
GSM_ORIG	161	179	-1.7	-0.11	$5.5 \times 10^{-20} *$
GSM_GC	220	204	-0.10	0.38	0.13
GSM_GS	221	203	-0.13	0.35	0.14
			SeaW	iFS	
OCx	183	124	-1.1	0.040	$2.1 \times 10^{-8} *$
POLY1	183	124	-0.32	0.29	$2.7 \times 10^{-3} *$
POLY4	183	124	-0.37	0.30	$5.8 \times 10^{-4} *$
GSM_ORIG	112	100	-1.7	-0.14	$3.8 \times 10^{-9} *$
GSM_GC	144	109	-0.17	0.32	$3.7 \times 10^{-3} *$
GSM_GS	144	109	-0.082	0.35	0.011 *
			VIII	RS	
OCx	51	121	-1.4	-0.061	9.0 × 10 ⁻⁷ *
POLY1	51	121	-0.42	0.16	0.083
POLY4	51	121	-0.48	0.17	0.038 *
GSM_ORIG	39	105	-1.2	-0.16	$1.0 \times 10^{-5} *$
GSM_GC	42	111	0.047	-0.036	0.87
GSM_GS	42	111	-0.098	0.11	0.33

Table A4. Results of the ANOVA on μ_{error} of seasonal subsets in the NEP, to check for statistically significant differences in means between seasons. The SeaWiFS subset has been excluded from this test, as all SeaWiFS matchups were collected during the summer, with the exception of a single matchup from fall and one from winter. Tukey Honest Significant Difference p-values between each pair of seasons are included to determine which seasons, if any, are statistically different from each other.

Algorithm	$N_{ m spring}$	N _{summer}	N_{fall}	$\mu_{error}^{ m spring}$	$\mu_{error}^{\mathrm{summer}}$	$\mu_{error}^{\mathrm{fall}}$	<i>p-</i> Value	$p_{ m spring}^{ m summer}$	$p_{ m spring}^{ m fall}$	p ^{fall} summer
					M	ODIS				
OCx	87	194	135	-0.28	0.23	-0.20	0.47	0.56	0.99	0.58
POLY1	87	194	135	-0.37	-0.054	-0.38	0.63	0.75	1.0	0.67
POLY4	87	194	135	-0.40	-0.14	-0.44	0.69	0.82	1.00	0.71
GSM_ORIG	47	139	58	-1.5	-0.89	-1.3	0.39	0.41	0.91	0.66
GSM_GC	65	155	97	0.35	0.13	0.065	0.89	0.92	0.89	0.99
GSM_GS	64	154	97	0.24	-0.020	-0.20	0.75	0.87	0.73	0.92
					v	IIRS				
OCx	67	115	116	-0.89	0.014	-0.51	0.34	0.34	0.83	0.60
POLY1	67	115	116	-0.70	0.17	-0.093	0.32	0.29	0.55	0.85
POLY4	67	115	116	-0.80	0.12	-0.11	0.23	0.21	0.41	0.87
GSM_ORIG	36	93	74	-1.5	-0.48	-1.4	0.085	0.19	0.98	0.14
GSM_GC	52	99	83	-0.52	0.30	0.29	0.42	0.44	0.47	1.0
GSM_GS	52	101	83	-0.68	0.97	-0.042	0.16	0.17	0.78	0.41

Remote Sens. 2019, 11, 2609 25 of 29

Table A5. Results of the *t*-test to check for differences in μ_{error} between coastal (bathymetry <= 200 m) and open ocean (>200 m) subsets. * = statistically significant (*p*-value < 0.05).

	N μ_{error}			Λ	I	μ				
Algorithm	Coastal	Open	Coastal	Open	<i>p</i> -Value	Coastal	Open	Coastal	Open	<i>p</i> -Value
			NWA		•			NE	P	•
					M	ODIS				
OCx	232	276	-0.27	-0.51	0.20	183	275	-0.46	0.18	0.064
POLY1	232	276	0.42	-0.20	$4.9 \times 10^{-3} *$	183	275	-0.78	0.083	$4.6 \times 10^{-3} *$
POLY4	232	276	0.32	-0.20	$6.7 \times 10^{-3} *$	183	275	-0.90	0.076	$8.4 \times 10^{-4} *$
GSM_ORIG	142	212	-1.0	-0.71	0.077	77	201	-2.4	-0.41	$7.3 \times 10^{-9} *$
GSM_GC	195	244	0.36	-0.066	0.17	118	237	0.17	0.12	0.91
GSM_GS	195	244	0.33	-0.090	0.19	118	236	-0.093	0.016	0.77
						aWiFS				
OCx	165	171	-0.53	-0.64	0.55	10	30	-0.17	-0.40	0.69
POLY1	165	171	0.19	-0.30	$6.9 \times 10^{-3} *$	10	30	0.36	-0.12	0.39
POLY4	165	171	0.14	-0.31	0.011 *	10	30	0.37	-0.13	0.35
GSM_ORIG	107	132	-1.1	-0.72	0.17	8	25	-1.1	-0.20	0.032 *
GSM_GC	134	148	0.24	-0.11	0.029 *	8	26	0.19	-0.047	0.44
GSM_GS	134	148	0.34	-0.091	$7.1 \times 10^{-3} *$	8	26	0.36	-0.12	0.17
			VIIRS							
OCx	74	98	-0.23	-0.61	0.13	135	195	-1.1	0.12	$4.5 \times 10^{-3} *$
POLY1	74	98	0.40	-0.32	0.021	135	195	-0.77	0.27	$9.0 \times 10^{-3} *$
POLY4	74	98	0.36	-0.31	0.020	135	195	-0.87	0.26	$2.4 \times 10^{-3} *$
GSM_ORIG	57	87	-0.47	-0.43	0.87	69	161	-2.3	-0.28	$5.5 \times 10^{-7} *$
GSM_GC	64	89	0.26	-0.21	0.29	88	177	0.29	$3.7 imes 10^{-4}$	0.56
GSM_GS	64	89	0.010	0.085	0.70	89	177	0.77	-0.069	0.20

Table A6. Average difference between first and third quartiles presented in the boxplots of Figures 6 and 7 across boxes representing the first half of the year, and across the second half of the year.

	POLY4	GSM_GS	POLY4	GSM_GS	POLY4	GSM_GS
	M	ODIS	Sea	aWiFS	V	IIRS
			N	IWA		
January–June	1.7	2.0	1.1	0.94	1.7	1.5
July-December	0.74	0.84	0.57	0.60	0.46	0.44
			1	NEP		
January–June	0.65	0.37	0.33	0.27	0.54	0.37
July-December	1.0	0.72	0.94	0.38	1.1	0.62

Table A7. Coefficients of the remaining OCx algorithms for each sensor and region.

	МО	DIS	Seal	WiFS	VIIRS		
Coefficients	POLY2	POLY3	POLY2	POLY3	POLY2	POLY3	
			NV	VA			
а	0.37539	0.37657	0.51424	0.52039	0.41461	0.44156	
b	-3.12409	-3.26173	-3.59265	-3.75269	-2.54637	-3.05795	
С	-0.75408	-0.60435	-0.95058	-0.92392	-1.47087	-0.65894	
d	-	1.1404	- 1.71524		-	1.21248	
e	-	-	-	-	-	-	
			N	EP			
а	0.28424	0.2805	0.42171	0.42506	0.33771	0.3303	
b	-2.66996	-2.77728	-2.95509	-2.74285	-2.56462	-2.74252	
С	-1.09915	-1.01747	-0.68104	-1.48743	-0.5314	-0.34545	
d	-	0.92282	-	0.17624	-	1.35569	
e	_	-	-	-	-	-	

Remote Sens. 2019, 11, 2609 26 of 29

Table A8. g coefficients used in Equation (11) after performing linear interpolation to the necessary wavelengths.

λ	<i>g</i> 1	<i>g</i> ₂	<i>g</i> ₃
400	0.0742	0.0805	1.4839
410	0.0716	0.0820	1.4520
420	0.0697	0.0841	1.4353
430	0.0685	0.0862	1.4300
440	0.0697	0.0890	1.4595
450	0.0773	0.1009	1.6387
460	0.0801	0.1142	1.7489
470	0.0832	0.1394	1.9055
480	0.0869	0.2095	2.1890
490	0.0878	0.2621	2.3091
500	0.0875	0.2820	2.3212
510	0.0861	0.2568	2.2215
520	0.0844	0.2233	2.1058
530	0.0821	0.1967	1.9920
540	0.0800	0.1811	1.9097
550	0.0781	0.1717	1.8464
560	0.0763	0.1651	1.7968
570	0.0754	0.1624	1.7722
580	0.0757	0.1640	1.7816
590	0.0768	0.1712	1.8211
600	0.0781	0.1864	1.8879
610	0.0784	0.1939	1.9143
620	0.0783	0.1956	1.9172
630	0.0782	0.1969	1.9186
640	0.0780	0.1973	1.9160
650	0.0782	0.2009	1.9283
660	0.0789	0.2227	1.9923
670	0.0798	0.2513	2.0663
680	0.0795	0.2465	2.0510
690	0.0789	0.2270	2.0000
700	0.0791	0.2323	2.0137

Table A9. Exponents used in Equations (6) and (10), where the g_c columns give exponents optimized to the original constant g_1 and g_2 coefficients, and g_s columns give the exponents optimized for the use of the spectral g coefficients g_1 , g_2 , and g_3 .

	MODIS		MODIS SeaWiFS VIIRS		MODIS S		Seal	SeaWiFS		VIIRS		
	NWA						NEP					
	g c	g_s	g c	g_s	g c	g_s	g c	g_s	g c	g_s	g_c	g_s
median P median S	0.500 0.038	0.500 0.036	0.500 0.035	0.500 0.034	0.600 0.026	0.500 0.026	0.600 0.038	0.600 0.036	0.700 0.028	0.650 0.026	0.600 0.034	0.600 0.030
median Y	0.800	0.750	0.600	0.525	1.400	1.750	0.900	0.750	0.750	0.650	0.800	0.750

Remote Sens. 2019, 11, 2609 27 of 29

Table A10. Spectrally-dependent a_{ph}^{\star} coefficients used in Equation (10) for the wavebands used in MODIS, SeaWiFS, and VIIRS.

λ	a_{ph}^{\star}
410	0.054343
412	0.055765
443	0.063252
469	0.051276
486	0.04165
488	0.040648
490	0.039546
510	0.025105
531	0.015745
547	0.011477
551	0.010425
555	0.009382
645	0.008967
667	0.019878
670	0.022861
671	0.023646
678	0.024389

References

- 1. Vargas, M.; Brown, C.W.; Sapiano, M.R.P. Phenology of marine phytoplankton from satellite ocean color measurements. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]
- 2. Racault, M.F.; Quéré, C.L.; Buitenhuis, E.; Sathyendranath, S.; Platt, T. Phytoplankton phenology in the global ocean. *Ecol. Indic.* **2012**, *14*, 152–163. [CrossRef]
- 3. Siegel, D.A.; Buesseler, K.O.; Doney, S.C.; Sailley, S.F.; Behrenfeld, M.J.; Boyd, P.W. Global assessment of ocean carbon export by combining satellite observations and food-web models. *Glob. Biogeochem. Cycles* **2014**, *28*, 181–196. [CrossRef]
- Toming, K.; Kutser, T.; Uiboupin, R.; Arikas, A.; Vahter, K.; Paavel, B. Mapping Water Quality Parameters with Sentinel-3 Ocean and Land Colour Instrument imagery in the Baltic Sea. *Remote Sens.* 2017, 9, 1070. [CrossRef]
- 5. Natvik, L.J.; Evensen, G. Assimilation of ocean colour data into a biochemical model of the North Atlantic: Part 1. Data assimilation experiments. *J. Mar. Syst.* **2003**, *40-41*, 127–153. [CrossRef]
- 6. IOCCG. *Remote Sensing in Fisheries and Aquaculture*; Reports of the International Ocean Colour Coordinating Group; IOCCG: Dartmouth, NS, Canada, 2009; Volume 8. [CrossRef]
- 7. McIver, R.; Breeze, H.; Devred, E. Satellite remote-sensing observations for definitions of areas for marine conservation: Case study of the Scotian Slope, Eastern Canada. *Remote Sens. Environ.* **2018**, 214. [CrossRef]
- 8. Johnson, C.; Devred, E.; Casault, B.; Head, E.; Cogswell, A.; Spry, J. Optical, Chemical, and Biological Oceanographic Conditions on the Scotian Shelf and in the Eastern Gulf of Maine during 2015. Available online: http://publications.gc.ca/site/eng/9.833512/publication.html (accessed on 5 November 2019).
- 9. Hannah, C.G.; McKinnell, S. (Eds.) *Applying Remote Sensing Data to Fisheries Management in BC*; Technical Report; Department of Fisheries and Oceans Canada: Ottawa, ON, Canada, 2016.
- 10. 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 algorithm for SeaWiFS. *J. Geophys. Res.* **1998**, 103, 24937–24953. [CrossRef]
- 11. Werdell, J.; Franz, B.; Bailey, S.; Feldman, G.; Boss, E.; Brando, V.; Dowell, M.; Hirata, T.; Lavender, S.; Lee, Z.; et al. Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* **2013**, *52*, 2019–2037. [CrossRef]
- 12. Werdell, P.J.; Bailey, S.W. An improved *in situ* bio-optical dataset for ocean colour algorithm development and satellite data production validation. *Remote Sens. Environ.* **2005**, *98*, 122–140. [CrossRef]
- 13. Gower, J. On the use of satellite-measured chlorophyll fluorescence for monitoring coastal waters. *Int. J. Remote Sens.* **2015**. [CrossRef]
- 14. Doerffer, R.; Schiller, H. The MERIS case 2 water algorithm. Int. J. Remote Sens. 2007, 28, 517–535. [CrossRef]

Remote Sens. 2019, 11, 2609 28 of 29

15. Laliberté, J.; Larouche, P.; Devred, E.; Craig, S. Chlorophyll-a Concentration Retrieval in the Optically Complex Waters of the St. Lawrence Estuary and Gulf Using Principal Component Analysis. *Remote Sens.* **2018**, *10*, 265. [CrossRef]

- 16. Hamed, G.; Scott, R. Revisiting empirical ocean-colour algorithms for remote estimation of chlorophyll-a content on a global scale. *Int. J. Remote Sens.* **2016**, *37*, 2682–2705. [CrossRef]
- 17. Ben Mustapha, S.; Bélanger, S.; Larouche, P. Evaluation of ocean color algorithms in the southeastern Beaufort Sea, Canadian Arctic: New parameterization using SeaWiFS, MODIS, and MERIS spectral bands. *Can. J. Remote Sens.* **2012**, *38*, 535–556. [CrossRef]
- 18. Peña, A.; Nemcek, N. Phytoplankton in Surface Waters along Line P and off the West Coast of Vancouver Island. In *State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2017*; Chandler, P.C., King, S.A., Boldt, J., Eds.; Can. Tech. Rep. Fish. Aquat. Sci.; 2018; Volume 3266, pp. 55–59. Available online: http://waves-vagues.dfo-mpo.gc.ca/Library/40717914.pdf (accessed on 5 November 2019).
- 19. Head, E.J.H.; Horne, E.P.W. Pigment transformation and vertical flux in an area of convergence in the North Atlantic. *Deep-Sea Res. II* **1993**, *40*, 329–346. [CrossRef]
- 20. Stuart, V.; Head, E.J.H. The BIO method. In *The Second SeaWiFS HPLC Analysis Round-Robin Experiment* (SeaHARRE-2); NASA/TM 2005-212785; Hooker, S.B., Ed.; NASA, Goddard Space Flight Center: Greenbelt, MD, USA, 2005; p. 112.
- 21. Zapata, M.; Rodríguez, F.; Garrido, J. Separation of chlorophylls and carotenoids from marine phytoplankton: A new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases. *Mar. Ecol. Prog. Ser.* **2000**, *195*, 29–45. [CrossRef]
- 22. Hijmans, R.J. Geosphere: Spherical Trigonometry (R Package Version 1.5-7). 2017. Available online: http://cran.nexr.com/web/packages/geosphere/index.html (accessed on 5 November 2019).
- 23. Maritorena, S.; Siegel, D.A.; Peterson, A.R. Optimization of a semi-analytical ocean color model for global-scale applications. *Appl. Opt.* **2002**, *41*, 2705–2713. [CrossRef]
- 24. R Core Team. *R: A Language and Environment for Statistical Computing;* R Foundation for Statistical Computing: Vienna, Austria, 2016.
- 25. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Applications*; Cambridge University Press: Cambridge, UK, 1997; ISBN 0-521-57391-2.
- 26. Canty, A.; Ripley, B.D. Boot: Bootstrap R (S-Plus) Functions (R Package Version 1.3-22). 2019. Available online: https://cran.rapporter.net/bin/linux/ubuntu/disco-cran35/Packages (accessed on 5 November 2019).
- 27. Garver, S.A.; Siegel, D.A. Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea. *J. Geophys. Res. Oceans* **1997**, *102*, 18607–18625. [CrossRef]
- 28. Gordon, H.R.; Brown, J.W.; Evans, R.H.; Brown, J.W.; Smith, R.C.; Baker, K.S.; Clark, D.K. A semianalytic radiance model of ocean color. *J. Geophys. Res.* **1988**, *93*, 10909–10924. [CrossRef]
- 29. Pope, R.M.; Fry, E.S. Absorption spectrum (380–700 nm) of pure water. II. Integrating measurements. *Appl. Opt.* **1997**, *36*, 8710–8723. [CrossRef]
- 30. Smith, R.; Baker, K. Optical properties of the clearest natural waters (200–800 nm). *Appl. Opt.* **1981**, 20, 177–184. [CrossRef] [PubMed]
- 31. IOCCG. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications; Reports of the International Oean-Colour Coordinating Group, No. 5; Lee, Z.P., Ed.; IOCCG: Dartmouth, NS, Canada, 2006.
- 32. Lee, Z.P.; Carder, K.L.; Arnone, R. Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters. *Appl. Opt.* **2002**, *41*, 5755–5772. [CrossRef] [PubMed]
- 33. Bricaud, A.; Claustre, H.; Oubelkheir, K. Natural variability of phytoplankton absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* **2004**, *110*. [CrossRef]
- 34. Legendre, P. Lmodel2: Model II Regression (R Package Version 1.7-3). 2018. Available online: https://cran.r-project.org/web/packages/lmodel2/index.html (accessed on 5 November 2019).
- 35. Seegers, B.; Stumpf, R.; Schaeffer, B.; Loftin, K.; Werdell, J. Performance metrics for the assessment of satellite data products: An ocean color case study. *Opt. Express* **2018**, *26*, 7404. [CrossRef] [PubMed]

Remote Sens. 2019, 11, 2609 29 of 29

36. Brewin, B.; Sathyendranath, S.; Müller, D.; Brockmann, C.; Deschamps, P.Y.; Devred, E.; Doerffer, R.; Fomferra, N.; Franz, B.; Grant, M.; et al. The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ.* 2015, 162. [CrossRef]

- 37. Wickham, H. ggplot2: Elegant Graphics for Data Analysis; Springer: New York, NY, USA, 2009.
- 38. Claustre, H. The trophic status of various oceanic provinces as revealed by phytoplankton pigment signatures. *Limnol. Oceanogr.* **1994**, *39*, 1206–1210. [CrossRef]
- 39. Jeffrey, S.W.; Vesk, M. Introduction to marine phytoplankton and their pigment signature. In *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*; Jeffrey, S.W., Mantoura, R.F.C., Wright, S.W., Eds.; UNESCO Publishing: Paris, France, 1997; pp. 37–84.
- 40. Garcia, V.M.T.; Signorini, S.; Garcia, C.A.E.; McClain, C.R. Empirical and semi-analytical chlorophyll algorithms in the south-western Atlantic coastal region (25–40°S and 60–45°W). *Int. J. Remote Sens.* **2006**, 27, 1539–1562. [CrossRef]
- 41. Sun, L.; Guo, M.; Wang, X. Ocean color products retrieval and validation around China coast with MODIS. *Acta Oceanol. Sin.* **2010**, 29, 21–27. [CrossRef]
- 42. Jiang, W.; Knight, B.R.; Cornelisen, C.; Barter, P.; Kudela, R. Simplifying Regional Tuning of MODIS Algorithms for Monitoring Chlorophyll-a in Coastal Waters. *Front. Mar. Sci.* **2017**, *4*, 151. [CrossRef]
- 43. Hooker, S.B.; Esaias, W.E.; Feldman, G.C.; Gregg, W.W.; McClain, C.R. *An Overview of SeaWiFS and Ocean Color*; Tech. Memo. 104566; Hooker, S., Firestone, E., Eds.; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 1992; Volume 1, 24p.
- 44. Eplee, R.E.; Robinson, W.D.; Bailey, S.W.; Clark, D.K.; Werdell, P.J.; Wang, M.; Barnes, R.A.; McClain, C.R. Calibration of SeaWiFS. II. Vicarious techniques. *Appl. Opt.* **2001**, *40*, 6701–6718. [CrossRef]
- 45. Stuart, V.; Sathyendranath, S.; Platt, T.; Maass, H.; Irwin, B.D. Pigments and species composition of natural phytoplankton populations: effect on the absorption spectra. *J. Plankton Res.* **1998**, 20, 187–217. [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. Devred, E.; Sathyendranath, S.; Stuart, V.; Maass, H.; Ulloa, O.; Platt, T. A two-component model of phytoplankton absorption in the open ocean: theory and applications. *J. Geophys. Res.* **2006**, *111*. [CrossRef]
- 48. Sathyendranath, S.; Watts, L.; Devred, E.; Platt, T.; Caverhill, C.; Maass, H. Discrimination of diatoms from other phytoplankton using ocean-colour data. *Mar. Ecol. Prog. Ser.* **2004**, 272, 59–68. [CrossRef]
- 49. IOCCG. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products; Reports of the International Ocean Colour Coordinating Group; IOCCG: Dartmouth, NS, Canada, 2010; Volume 10. [CrossRef]
- 50. Morel, A.; Prieur, L. Analysis of variation in ocean color. Limnol. Oceanogr. 1977, 22, 709–722. [CrossRef]
- 51. Carswell, T.; Costa, M.; Young, E.; Komick, N.; Gower, J.; Sweeting, R. Evaluation of MODIS-Aqua Atmospheric Correction and Chlorophyll Products of Western North American Coastal Waters Based on 13 Years of Data. *Remote Sens.* 2017, *9*, 1063. [CrossRef]



© 2019 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/).