Revisiting the Ocean Color Algorithms for Particulate Organic Carbon and Chlorophyll-a Concentrations in the Ross Sea

Shuangling Chen
Walker O. Smith Jr.
Xiaolei Yu

Follow this and additional works at: https://scholarworks.wm.edu/vimsarticles

Part of the Oceanography Commons
Revisiting the Ocean Color Algorithms for Particulate Organic Carbon and Chlorophyll-a Concentrations in the Ross Sea

Shuangling Chen1, Walker O. Smith Jr.2,3, and Xiaolei Yu1,2

1State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China, 2School of Oceanography, Shanghai Jiao Tong University, Shanghai, China, 3Virginia Institute of Marine Science, Gloucester Point, VA, USA

Abstract The Ross Sea is the most productive marginal sea in the Southern Ocean and plays an important role in carbon cycling. However, limited sampling of Chlorophyll-a (Chl) and particulate organic carbon (POC) concentrations from research expeditions constrains our understanding of the biogeochemical processes there. Satellites provide a useful tool for synoptic mapping of surface water properties on regional and global scales, yet the general applicability of the published algorithms in the Ross Sea is poorly known. Based on the data collected from 18 cruises in the past 20 years, we analyzed both the NASA standard and locally developed Chl and POC algorithms applicable to the Ross Sea. Our results show that Chl and POC are markedly underestimated using the NASA standard algorithms, with root mean square difference (RMSD) of 4.72 mg m⁻³ and 218.0 mg m⁻³, and mean bias of −3.48 mg m⁻³ and −159.1 mg m⁻³, for a wide range of Chl (0.42–16.3 mg m⁻³) and POC (46.8–812 mg m⁻³). Similar poor performances were also found for other algorithms applicable in the Ross Sea. We locally tuned both Chl and POC algorithms, and found that the Rrs667-based approach showed the most robust performances in retrieving both Chl and POC, with improved RMSD of 2.86 mg m⁻³ and 129.7 mg m⁻³, and limited biases. Our results show that the algal bloom signals in the Ross Sea in terms of Chl and POC are significantly greater than previously determined. More field observations will further constrain the locally tuned algorithms.

Plain Language Summary The Ross Sea is the most productive marginal sea in the Southern Ocean and plays an important role in carbon cycling. Chlorophyll-a (Chl) and particulate organic carbon (POC) concentrations are important proxies of phytoplankton biomass in the ocean. However, sampling of Chl and POC is very limited from research expeditions there. Satellites provide a useful tool for synoptic mapping of surface water properties on regional and global scales, yet the general applicability of the published Chl and POC algorithms in the Ross Sea are poorly known. Based on the data collected in the past 20 years, we revisited all the published Chl and POC algorithms applicable to the Ross Sea. Our results show that Chl and POC are markedly underestimated or overestimated using these algorithms. We developed both Chl and POC algorithms based on ocean color and found that the Rrs667-based approach showed the most robust performances in retrieving both Chl and POC, and the uncertainties are greatly reduced with small biases. Using these updated algorithms, we found that the algal bloom signals in the Ross Sea (Chl and POC) are significantly higher than previously determined. More field observations will further constrain the locally tuned algorithms.

1. Introduction

The ocean plays an important role in the global carbon cycle, absorbing ~25% of anthropogenic CO₂ since the global industrialization (Friedlingstein et al., 2019; Gruber et al., 2019; Sabine et al., 2004). Phytoplankton photosynthesis in the ocean’s surface layers utilize nutrients and convert dissolved CO₂ to particulate organic carbon (POC) and cellular components. Chlorophyll-a (Chl) and POC are important parameters in biogeochemical studies, particularly the biological carbon pump, which is one of the significant mechanisms to transport carbon from surface to the deep ocean.

Ocean color remote sensing is an effective tool for understanding ocean ecology and biogeochemistry on synoptic scales. Since the launch of the Coastal Zone Color Scanner (1978–1986), ocean color has been...
widely used to derive surface biogeochemical parameters, particularly Chl and POC from spectral remote sensing reflectance \([Rrs(\lambda)]\) throughout the ocean (Hu et al., 2012; Le et al., 2018; O'Reilly et al., 2000; Stramski et al., 2008). Tremendous success has been achieved in the past decades in terms of long-term data records of several key biogeochemical parameters. Special effort has been placed on the retrieval accuracies of ocean color data products. For example, the Chl algorithm has been updated from \(Rrs(\lambda)\) band-ratio-based OCx versions to \(Rrs(\lambda)\) band-difference-based OCI version to improve the algorithm's accuracy (Hu et al., 2012; O'Reilly et al., 2000). Several studies have estimated POC from optical backscattering (Behrenfeld et al., 2008; Stramski et al., 1999), yet the relationship appears to vary spatially in waters even with similar levels of POC (Stramski et al., 1999). Pabi and Arrigo (2006) suggested that some of this variability could be attributed to the spatial differences in the taxonomic composition of the phytoplankton and their associated microbial communities. Taking advantage of the linkage between \(Rrs(\lambda)\) and particulate backscattering, Stramski et al. (2008) developed a powerful relationship between POC and \(Rrs(\lambda)\) blue-green band ratios \((R_{bg} = Rrs443/Rrs555)\) in the open ocean waters; this algorithm is now implemented in the NASA standard data processing routine. Overall, the present NASA standard Chl and POC algorithms are reported to have an uncertainty of 16.5%–30% and 20%, respectively, on global scales (Hu et al., 2012; Stramski et al., 2008), which are within the operational goal of 35%.

Accuracy is critical to ocean retrieval because time-series analyses call for the most robust products to study temporal changes. However, the present standard Chl and POC algorithms in marginal seas tend to have larger uncertainties because of optical complexities in coastal waters. For this reason, many regional algorithms of Chl are developed based on the optical characteristics in a specific coastal environment (e.g., Gohin et al., 2002; Moses et al., 2012; Pahlevan et al., 2020; Vilas et al., 2011; Zheng & DiGiacomo, 2017). Similarly, the present standard Chl and POC algorithms may not be suitable in remote polar regions, where \textit{in situ} data are limited.

The polar regions exhibit bio-optical properties that differ significantly from the mid- and low-latitude oceans (Szeto et al., 2011). Cloud cover is persistent, particularly over open ocean waters where substantial heat exchange occurs, and solar angles are often low, even in summer. To acclimate to low light conditions, polar phytoplankton package pigments to increase the total light absorbed per cell compared to phytoplankton in lower latitude waters, resulting in a lower light absorption per Chl and an underestimation of Chl using the standard Chl algorithm (Matsuoka et al., 2007, 2011; Mitchell, 1992). Lewis et al. (2016) evaluated the global Chl algorithm in the Arctic Ocean, and found that Chl was overestimated when Chl < 0.9 mg m\(^{-3}\) due to CDOM contamination and was underestimated at a higher concentration because of pigment packaging. Haëntjens et al. (2017) revisited the standard Chl and POC algorithms for the Southern Ocean open waters using field observations from biogeochemical floats, and they found good agreement between the satellite-derived Chl and POC and float-observed Chl and POC. However, for the marginal seas around Antarctica, several studies have found that the NASA standard algorithm underestimates Chl (Dierssen & Smith, 2000; Garcia et al., 2005; Mitchell & Holm-Hansen, 1991; Mitchell et al., 2001). For the eastern Antarctic, particularly the Ross Sea, the general applicability of the standard Chl and POC algorithms is poorly known. Pabi and Arrigo (2006) tried to develop regional surface POC algorithms for SeaWiFS with a linear fit to \(Rrs(555\text{ nm})\) for each phytoplankton taxa (\textit{Phaeocystis antarctica} and diatom) for the Ross Sea based on surface measurements (i.e., \(Rrs\) spectra and POC) from two cruises in 1996–1998, yet the difficulty of remotely discriminating phytoplankton taxa definitely limits its general application on satellite images.

The continental shelf of the Ross Sea is known to be the most productive marginal sea in the Southern Ocean particularly in spring and summer when the coastal polynyas form and phytoplankton blooms, with a mean productivity of 0.050 Pg C yr\(^{-1}\) (Arrigo, van Dijken, & Bushinsky, 2008), and it also serves as a strong CO\(_2\) sink of 0.013 Pg C yr\(^{-1}\) (Arrigo, van Dijken, & Long, 2008). There are two dominant functional groups in the region, the haptophyte \textit{P. antarctica} and diatoms (Jones & Smith, 2017; Smith et al., 2014). Based on 42 cruises on the Ross Sea continental shelf accumulated in the past decades, Smith and Kaufman (2018) found that the POC covaried with Chl through spring and early summer, but there was a notable increase in POC in later summer that was largely independent to Chl changes. They speculated that this change resulted from an adaptation of diatoms to extreme iron limitation. Better quantification of Chl and POC from satellite images would improve our ability to examine the spatial and temporal dynamics of POC and Chl, which would then promote our understanding of the biogeochemical cycling in this important region.
To facilitate a better understanding of the synoptic distributions of POC and Chl, we analyzed the remotely sensed Chl and POC and compared those to data obtained during 18 cruises that measured surface Chl and POC measurements simultaneously since 1997 (when the ocean color satellite missions began). Our objectives of this study included: (a) evaluating the performance of NASA's standard and other regional algorithms of Chl and POC in the Ross Sea; (b) improving the remote estimations of Chl and POC; and (c) analyzing the spatial and temporal dynamics of Chl and POC on synoptic scales.

### 2. Data and Methods

#### 2.1. Data

Cruise data of POC and Chl were synthesized from a total of 18 cruises (Table 1) and 605 stations in the Ross Sea (Figure 1) between 1997 and 2013. The compiled data at all depths were obtained from [https://scholar-works.wm.edu/data/380](https://scholar-works.wm.edu/data/380) (Smith & Kaufman, 2018). Most cruises occurred in late spring and summer after the formation of polynya with the meltwater from sea ice on the continental shelf, and only one cruise (i.e., cruise "PRISM") had data samples in the open waters of the Ross Sea. The vertical sampling frequency varied among these cruises, with a few samples collected in the upper 10 meters and every 10 m or 20 m below. Only the surface POC and Chl data (i.e., averages of the upper 5 m layer, within a standard deviation of <8%) from these cruises were used in this study. Both surface Chl and POC concentrations show large variations within each cruise (see Table 1), and Chl and POC varied within 0.29–0.99 mg m$^{-3}$ and 21.6–88.08 mg m$^{-3}$, respectively, in open waters. The Chl concentrations were usually determined either by standard fluorometric methods (JGOFS, 1996), but occasionally by high performance liquid chromatography (DiTullio et al., 2003). Previous studies showed no significant difference between these two methods in the Ross Sea (Smith et al., 2000). POC concentrations were determined by high temperature pyrolysis on elemental

#### Table 1

<table>
<thead>
<tr>
<th>Cruise</th>
<th>Date</th>
<th>Chl range (mg m$^{-3}$)</th>
<th>POC range (mg m$^{-3}$)</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVARS_I_Leg_1</td>
<td>12/19/2001–12/21/2001</td>
<td>2.64–10.3</td>
<td>173.3–470.8</td>
<td>11</td>
</tr>
<tr>
<td>IVARS_I_Leg_2</td>
<td>02/03/2002–02/07/2002</td>
<td>2.32–7.69</td>
<td>72.4–515.0</td>
<td>17</td>
</tr>
<tr>
<td>IVARS_II_Leg_1</td>
<td>12/23/2002–12/28/2002</td>
<td>0.35–6.7</td>
<td>NaN</td>
<td>10</td>
</tr>
<tr>
<td>IVARS_II_Leg_2</td>
<td>02/10/2003</td>
<td>2.20–3.48</td>
<td>NaN</td>
<td>2</td>
</tr>
<tr>
<td>IVARS_III_Leg_1</td>
<td>12/26/2003–12/30/2003</td>
<td>1.43–9.48</td>
<td>142.0–509.7</td>
<td>16</td>
</tr>
<tr>
<td>IVARS_III_Leg_II</td>
<td>02/03/2004–02/09/2004</td>
<td>4.14–14.4</td>
<td>216.0–516.3</td>
<td>25</td>
</tr>
<tr>
<td>IVARS_IV_Leg_1</td>
<td>12/19/2004–12/24/2004</td>
<td>0.86–13.4</td>
<td>156.9–847.4</td>
<td>23</td>
</tr>
<tr>
<td>IVARS_IV_Leg_2</td>
<td>01/29/2005–01/31/2005</td>
<td>0.16–3.06</td>
<td>81.7–255.7</td>
<td>22</td>
</tr>
<tr>
<td>CORSACS_YR_1</td>
<td>12/27/2005–01/23/2006</td>
<td>0.69–5.45</td>
<td>43.5–706.0</td>
<td>60</td>
</tr>
<tr>
<td>IVARS_V_Leg_1</td>
<td>12/28/2005–12/30/2005</td>
<td>2.33–4.36</td>
<td>375.8–577.3</td>
<td>12</td>
</tr>
<tr>
<td>IVARS_V_Leg_2</td>
<td>01/31/2006–02/01/2006</td>
<td>0.48–2.14</td>
<td>20.0–117.1</td>
<td>12</td>
</tr>
<tr>
<td>CORSACS_YR2</td>
<td>11/14/2006–12/01/2006</td>
<td>0.28–7.11</td>
<td>26.7–331.7</td>
<td>42</td>
</tr>
<tr>
<td>OREN_2007</td>
<td>12/27/2007–01/02/2008</td>
<td>0.29–8.37</td>
<td>50.5–563.5</td>
<td>9</td>
</tr>
<tr>
<td>NBP11 - 01_POC</td>
<td>01/20/2011–02/08/2011</td>
<td>0.37–6.23</td>
<td>64.2–542.2</td>
<td>42</td>
</tr>
<tr>
<td>OREN_2011 - 12</td>
<td>12/31/2011</td>
<td>3.29</td>
<td>73.1–73.1</td>
<td>1</td>
</tr>
<tr>
<td>PRISM</td>
<td>01/10/2012–02/06/2012</td>
<td>0.29–22.4</td>
<td>19.7–999.1</td>
<td>114</td>
</tr>
<tr>
<td>NBP13 - 02</td>
<td>02/12/2013–03/09/2013</td>
<td>0.14–8.94</td>
<td>20.3–821.7</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>11/15/1997–03/09/2013</td>
<td>0.14–22.4</td>
<td>19.7–1111.7</td>
<td>605</td>
</tr>
</tbody>
</table>

Note. NaN = no data available. Cruises denoted in italics were used as independent data set for model validation, and the additional cruises were used for local model training.

To facilitate a better understanding of the synoptic distributions of POC and Chl, we analyzed the remotely sensed Chl and POC and compared those to data obtained during 18 cruises that measured surface Chl and POC measurements simultaneously since 1997 (when the ocean color satellite missions began). Our objectives of this study included: (a) evaluating the performance of NASA's standard and other regional algorithms of Chl and POC in the Ross Sea; (b) improving the remote estimations of Chl and POC; and (c) analyzing the spatial and temporal dynamics of Chl and POC on synoptic scales.
analyzers (Gardner et al., 2000). Dominant functional groups were determined either microscopically or by using CHEMTAX (DiTullio et al., 2003; Smith et al., 2010).

Instead of using weighted averages, we chose to use the *in situ* surface Chl and POC measurements to evaluate the satellite-derived Chl (Chl$^{RS}$) and POC (POC$^{RS}$). Because the satellite Rrs ($\lambda$) has information of Chl at depth in the upper water column, it was then argued that the Chl$^{RS}$ should be compared with the weighted average of Chl. However, the weighted Chl is both wavelength-dependent and weighting-formula-dependent, and there is no consensus on which weighting schemes should be used to compare with Chl$^{RS}$ (André, 1992; Morel & Berthon, 1989; Sathyendranath & Platt, 1989; Sokolovsky & Yacobi, 2011; Werdell & Bailey, 2005). Using numerically simulated data, André (1992) compared weighted Chl at 520 nm with Chl$^{RS}$ derived from Rrs(443)/Rrs(550) band ratio (i.e., $R_{bg}$), and concluded that the weighted average is necessary only under conditions of strong stratification near the surface. Lee et al. (2020) extended this analysis to other wavelengths along with much stronger stratification of Chl, and concluded that weighted averages of stratified waters are not required for empirical algorithms. In addition, Lee et al. (2020) and Sathyendranath et al. (2019) both emphasized that surface Chl is easily collected and is generally available for all field measurements, making it ideal to use surface Chl for the development and subsequent evaluation of the empirical algorithms. More importantly, it bypasses the requirement to have profiles of optical properties for the calculation of weighted averages. Therefore, using surface data provides an excellent option for algorithm development and verification.

The NASA standard daily Level-3 data products (version R2018.0) derived from measurements by the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) on the SeaStar satellite and the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua satellite were used in this study. Data of spectral Rrs ($\lambda$) at bands of 412, 443, 488 (490 nm for SeaWiFS), 547 (555 nm for SeaWiFS), and 667 nm (670 nm for SeaWiFS), as well as the NASA standard data products of Chl and POC, covering the study region from 1997 to 2020 with a spatial resolution of ~4 km, were downloaded from the NASA Goddard Space Flight Center
The MODIS-derived Rrs at 547 nm was converted to Rrs555 following Equation 1 (Chen & Hu, 2017) for band consistency between sensors.

\[ R_{rs547} \leq sw \Rightarrow R_{rs555} = 10^{a_1 \times \log_{10} R_{rs547} - b_1} \] (1a)

where \( sw = 0.001723 \), \( a_1 = 0.986 \), and \( b_1 = 0.081495 \).

\[ R_{rs555} \geq sw \Rightarrow R_{rs555} = a_2 \times R_{rs547} - b_2 \] (1b)

where \( sw = 0.001723 \), \( a_2 = 1.03 \), and \( b_2 = 0.000216 \).

### 2.2. Methods

The cruise data were divided into two groups with one for model training, and the other for independent model validation (see Table 1). To obtain high quality satellite data, the time and location of the field Chl and POC measurements (Figures 1a and 1b, Table 1) were used to find the concurrent and co-located spectral Rrs (\( \lambda \)), Chl, and POC from SeaWiFS and MODIS using the following criteria: for individual field observations, a time window of the same day between the field and satellite measurements was used to extract the valid satellite data within a 4 × 4 km pixel centered at each sampling location of in situ Chl and POC observations. These matched data were used to represent the corresponding satellite measurements at each sampling location. As a result, a total number of 112 conjugate valid observations of field Chl and POC and associated satellite-derived Chl and POC and Rrs (\( \lambda \)) spectra were determined between 1997 and 2013 (Figures 1c, 1d, and 2). This conjugate data set was representative of the study area, with the observed Chl concentrations ranging between 0.42 and 16.3 mg m\(^{-3}\), and the observed POC ranging between 46.8 and 812 mg m\(^{-3}\). In this data set, 80 data points were for model training, and the rest for model validation. The Rrs spectra of each dominant phytoplankton type (i.e., *P. antarctica* or diatoms) varied greatly for a wide range of Chl and POC (Figure 2), typically with a reflectance peak at 555 nm at high Chl concentrations.
and low values at blue band of 443 nm due to strong phytoplankton absorption. Between the two phytoplankton functional groups, the Rrs spectra showed some different but indistinct spectral characteristics. Specifically, the P. antarctica population tended to show lower (higher) reflectance than the diatom in the blue (green) band of 443 (555) nm, with a smaller difference at the red band (Stuart et al., 2000). Regardless, the indiscriminate spectral characteristics suggest the difficulty in differentiating the two from space and emphasize the potential problems if Chl and POC algorithms are developed for each species.

A variety of algorithms applicable to the Ross Sea have been developed to estimate Chl and POC (Table 2). The current NASA standard Chl algorithm is a hybrid of the OCx and OCI algorithms. For Chl retrievals ≥0.2 mg m\(^{-3}\) the OCx algorithm is used, and for Chl retrievals ≤0.15 mg m\(^{-3}\) the OCI algorithm is used; between these values, the OCx and OCI algorithms are blended using a weighted approach (Hu et al., 2012). The OCx algorithm is based on the maximum band ratio index (\(R_{bgm}\), Equation 2), and the OCI algorithm is based on the color index (CI), a weighted relative band difference between Rrs555 and a baseline formed linearly between Rrs443 and Rrs670 (Equation 3, Hu et al., 2012). We named the CI here as \(C_{CI}\). Similar to the OCx Chl algorithm, the current NASA standard POC algorithm is based on the band ratio of Rrs443/Rrs555. It was originally developed for SeaWiFS, and for MODIS applications, NASA uses a band shift to obtain Rrs555 from Rrs443 (Equation 1). This algorithm was formulated with concurrent in situ measurements of both POC and Rrs (\(\lambda\)) (Stramski et al., 2008). The CI approach was recently used to remotely estimate POC for the global ocean (Le et al., 2018), but with a baseline formed between Rrs490 and Rrs670. To differentiate it from \(C_{CI}\), we named it as \(C_{CI}\) (Equation 4). This \(C_{CI}\)-based POC algorithm was parameterized with two components for \(C_{CI} ≥ −0.0005\ \text{sr}^{-1}\) and \(C_{CI} ≥ −0.0005\ \text{sr}^{-1}\), respectively (see Table 2). It is notable that this POC algorithm was constructed based on in situ POC measurements within the surface layer (<10 m) and concurrent SeaWiFS-derived Rrs (\(\lambda\)) (Le et al., 2018).

\[
\text{Chl}_{OCx} = 10^{a_0 + a_1 \times R_{bgm} + a_2 \times R_{bgm}^2 + a_3 \times R_{bgm}^3 + a_4 \times R_{bgm}^4}
\]  

\(2a\)

where \(a_0 = 0.3272\), \(a_1 = −2.9940\), \(a_2 = 2.7218\), \(a_3 = −1.2259\), and \(a_4 = −0.5683\) for SeaWiFS, implemented for Chl ≥ 0.2 mg m\(^{-3}\). For the OC3 algorithm applied to MODIS, \(R_{bgm}\) is defined against \(\max(\text{Rrs443, Rrs490, Rrs510})/\text{Rrs547}\), with regression coefficients adjusted.

\[
R_{bgm} = \log_{10} \left( \max \left( \text{Rrs443, Rrs490, Rrs510} \right) / \text{Rrs555} \right)
\]  

\(2b\)

The 443, 490, 510, and 555 nm wavelengths are used for consistency with SeaWiFS wavebands.
\[ \text{Chl}_{\text{OCI}} = 10^{b_0 + b_1 \cdot \text{Chl}_{\text{Hu}}} \]  

where \( b_0 = -0.4909, b_1 = 191.6590 \), implemented for \( \text{Chl} \leq 0.15 \text{ mg m}^{-3} \).

\[ \text{CI}_{\text{Hu}} = \text{Rrs}_{555} - \left( \text{Rrs}_{443} + (\text{Rrs}_{670} - \text{Rrs}_{443}) \times \frac{555 - 443}{670 - 443} \right) \]  

The wavelengths of 443, 555, and 670 nm are specified for SeaWiFS, and for MODIS they are 443, 547, and 667 nm.

\[ \text{CI}_{\text{Le}} = \text{Rrs}_{555} - \left( \text{Rrs}_{490} + (\text{Rrs}_{670} - \text{Rrs}_{490}) \times \frac{555 - 490}{670 - 490} \right) \]  

Again, the wavelengths of 490, 555, and 670 nm are for SeaWiFS, and for MODIS they are 488, 547, and 667 nm.

Several studies have shown that the bio-optical properties of phytoplankton from polar waters are significantly different from those in lower latitudes, and suggested that different bio-optical algorithms should be used for polar regions (Cota et al., 2004; Jena, 2017; Kerkar et al., 2021; Lewis et al., 2019; Mitchell, 1992; Stuart et al., 2000). At present, there are no locally tuned Chl algorithms for the Ross Sea, but several regional POC algorithms that include the Ross Sea have been developed. For example, based on in situ data of Rrs555 and POC, Pabi and Arrigo (2006) developed a POC algorithm for the Ross Sea, while Allison et al. (2010) developed a POC algorithm for the entire Southern Ocean based on in situ measurements of POC and Rrs(\( \lambda \)) and using the band ratio of Rrs443/Rrs555 (R_{bg}). In general, these algorithms were developed with different bands and different coefficients based on different data sources. We re-evaluated all of these published algorithms.

To assess the performance of each algorithm listed in Table 2, we calculated Chl or POC following the regression formula of each with its original parameterization (i.e., original regression coefficients) using the same conjugated data set (Figures 1c and 1d). Statistical measures including root mean square difference (RMSD), coefficient of determination (R^2), mean bias (MB, Equation 5), mean ratio (MR, Equation 6), unbiased percent difference (UPD, Equation 7), and mean relative difference (MRD, Equation 8) (Barnes & Hu, 2014; Chen et al., 2019), as well as median bias and median ratio were used to quantify the accuracy of the estimated Chl or POC from each algorithm.

\[ \text{MB} = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{y}_i - y_i \right) \]  

where \( \hat{y}_i \) is derived from the satellite data, and \( y_i \) is derived from the field data. For median bias, the medians instead of the means were calculated.

\[ \text{MR} = \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{y}_i}{y_i} \]  

\[ \text{UPD} (\%) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\hat{y}_i - y_i}{\hat{y}_i + y_i} \right) \times 100 \]  

\[ \text{MRD} (\%) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\hat{y}_i - y_i}{y_i} \right) \times 100 \]  

Following the regression formula listed, we also locally fitted the Rrs spectra with Chl and POC using the same training data set. In addition, other regression formula for Chl and POC were also tested based on the correlations between the Rrs spectra and Chl and POC. All the locally tuned models were further validated using the independent validation data set (Table 1). The modeled Chl and POC in each case were compared with the field Chl and POC with statistics of RMSD, R^2, MB, MR, UPD, and MRD. In addition to these statistics, the selected algorithms in retrieving Chl and POC in the Ross Sea were determined using pairwise comparison (i.e., a winner score [%]; Seegers et al., 2018) among all the tested models, based on
the Chl/POC residuals between modeled and observed Chl/POC; then the selected algorithms were applied to the MODIS daily images to calculate the satellite Chl and POC for the Ross Sea between 2002 and 2020.

3. Results

3.1. Evaluation of the NASA Standard Algorithms

We found that both Chl and POC are strongly underestimated with large uncertainties using these standard algorithms (Figure 3). All the matched data pairs are located on the continental shelf of the Ross Sea (Figures 1c and 1d). Statistically, the comparison between satellite-derived Chl and field Chl showed an RMSD of 4.72 mg m$^{-3}$, $R^2$ of 0.28, MB of −3.48 mg m$^{-3}$, MR of 0.40, UPD of 99.9%, and MRD of −60.1% for observed Chl concentrations ranging between 0.42 and 16.3 mg m$^{-3}$ (Table 2). Similarly, the satellite-derived POC had an RMSD of 218 mg m$^{-3}$, $R^2$ of 0.32, MB of −159 mg m$^{-3}$, MR of 0.66, UPD of 64.2%, and MRD of −33.9% for measured POC ranging between 46.8 and 812 mg m$^{-3}$ (Table 2). In the open ocean waters of the Ross Sea, based on the limited data sampling from cruise “PRISM” (Table 1), we also found some underestimation in Chl and POC using the NASA standard algorithms (Figure S1).

3.2. Evaluation of the Locally Tuned Algorithms for Ross Sea

The performances of the other three POC algorithms (see Section 2.2) applicable to the Ross Sea were also evaluated using their original model parameterizations and the corresponding derived statistics (Table 2). Among these algorithms including the NASA standard POC algorithm, the CI$_{Le}$-based POC algorithms tended to show slightly better performance than others based on the statistical measures. Specifically, the POC algorithm by Allison et al. (2010) developed for the Southern Ocean was established using the same regression formula as the NASA standard POC algorithm, but with different regression coefficients, and the derived statistics were closest to that of the standard algorithm. The POC algorithm by Le et al. (2018) for the global ocean showed better performance in most statistical measures except RMSD. Surprisingly, the POC algorithm by Pabi and Arrigo (2006) developed particularly for the Ross Sea showed the worst performance among all in terms of each statistical measure, with an RMSD of 834.7 mg m$^{-3}$ and an MB of 655.7 mg m$^{-3}$ (Median bias = 752.5 mg m$^{-3}$).

The relationships between field-measured POC (Chl) and different Rrs($\lambda$) formulas available in the literature (Table 2), including Rrs555, $R_{bg}$, $R_{bgm}$, and CI (both CI$_{Hu}$ and CI$_{Le}$), were compared (Figures 4 and 5). Despite the poor performance of each algorithm (Table 2), we found significant correlations for each pair of variables examined, with correlation coefficients $R = −0.56$–$0.70$ ($p < 0.05$). Additionally, we found the largest correlations of POC and Chl with Rrs667, with $R$ values of 0.73 ($p < 0.05$).
Based on the correlations (Figure 4), the performance of each of the locally tuned POC algorithms was tested using the training and validation data set (Figure 6), with statistics summarized in Table 3. All these algorithms showed better performance than those with their original parameterizations (Table 2), with RMSD of 129.7–153.0 mg m\(^{-3}\). To a different extent, all the algorithms overestimate and underestimate POC for POC < ~400 mg m\(^{-3}\) and POC > ~400 mg m\(^{-3}\), respectively. Among these locally tuned POC algorithms, we found that the POC algorithm based on Rrs667 showed the most robust model performance (RMSD = 129.7 mg m\(^{-3}\), \(R^2 = 0.46\), MB = 2.26 mg m\(^{-3}\), MR = 1.2), and the POC algorithm using CI\(_{Le}\) showed a slightly poorer performance than the Rrs667-based POC algorithm.

Similarly, based on the correlations in Figure 5, we regressed Chl against the different Rrs(\(\lambda\)) formulas using the training data set, and validated it using the independent validation data set. The accuracies of the modeled Chl from each of these locally tuned algorithms were also greatly improved compared to that of the
NASA standard Chl algorithm, with an RMSD of 2.86–3.39 mg m$^{-3}$ (Figure 7; see more statistics in Table 4). Again, to a different extent, these algorithms show an overestimation and underestimation for Chl $<\sim 5$ mg m$^{-3}$ and Chl $>\sim 5$ mg m$^{-3}$, respectively. Among these locally tuned Chl algorithms, the Rrs667-based Chl algorithm had the most robust model performance (RMSD = 2.86 mg m$^{-3}$, $R^2 = 0.42$, MB = 0.38 mg m$^{-3}$, MR = 1.17); as before, the Chl algorithm using CI$_{Le}$ showed a slightly poorer performance than the one using Rrs667.

### 3.3. Chl and POC Dynamics

The daily Chl and POC images on January 3, 2014, using either the NASA default processing algorithms or the locally tuned Rrs667-based algorithms, were derived. From the MODIS default images (Figures 8a...
Figure 6. Performances of different particulate organic carbon (POC) algorithms in retrieving POC using the same training data set (blue dots) and validation data set (green dots), based on (a) Rrs555, (b) Rrs667, linear fitting of $R_{\text{bg}}$ (linear fitting in (c) and power fitting in (d)), $R_{\text{bgm}}$ (linear fitting in (e) and OCx fitting formula in (f)), (g) CI$_{Hu}$, and (h) CI$_{Le}$. The red lines indicate a 1:1 relationship. The statistics shown were based on the whole conjugate data set including both the training data set and the validation data set. The red line is the 1:1 line, the solid blue line is the fitting curve, with ± one standard deviation (blue dashed lines). See Table 3 for more detailed statistics.
Table 3
Model Comparisons of the Locally Tuned Particulate Organic Carbon (POC) Algorithms in the Ross Sea

<table>
<thead>
<tr>
<th>Regress formula</th>
<th>N</th>
<th>RMSD (mg m$^{-3}$)</th>
<th>$R^2$</th>
<th>MB</th>
<th>Median bias</th>
<th>MR</th>
<th>Median ratio</th>
<th>UPD (%)</th>
<th>MRD (%)</th>
<th>Fitting curve</th>
<th>Winner%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{a}$POC = $a$ * log$<em>{10}$(R$</em>{bg}$) + $b$</td>
<td>71/29</td>
<td>157.9/105.8</td>
<td>0.30/0.31</td>
<td>0.00/-23.13</td>
<td>67.93/-8.72</td>
<td>1.38/1.03</td>
<td>1.01/0.95</td>
<td>41.01/23.78</td>
<td>37.71/2.82</td>
<td>$y = 0.30x + 223.26$</td>
<td>2.50/9.38</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>144.7</td>
<td>0.30</td>
<td>-6.71</td>
<td>-7.30</td>
<td>1.28</td>
<td>0.98</td>
<td>36.01</td>
<td>27.59</td>
<td>-</td>
<td>11.88</td>
</tr>
<tr>
<td>$^{b}$POC = $a$ * (R$_{bg}$)$^b$</td>
<td>71/29</td>
<td>154.7/111.0</td>
<td>0.33/0.30</td>
<td>-1.38/-31.04</td>
<td>91.83/-24.84</td>
<td>1.35/1.00</td>
<td>1.02/0.93</td>
<td>40.48/25.12</td>
<td>35.31/-0.48</td>
<td>$y = 0.34x + 209.05$</td>
<td>8.75/21.88</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>134.3</td>
<td>0.32</td>
<td>-9.98</td>
<td>0.51</td>
<td>1.25</td>
<td>1.00</td>
<td>36.03</td>
<td>24.93</td>
<td>-</td>
<td>30.63</td>
</tr>
<tr>
<td>$^{c}$POC = $a$ * log$<em>{10}$(R$</em>{bgm}$) + $b$</td>
<td>69/22</td>
<td>158.8/95.0</td>
<td>0.30/0.51</td>
<td>0.00/-22.50</td>
<td>71.91/-16.52</td>
<td>1.37/1.03</td>
<td>1.00/0.96</td>
<td>40.15/20.22</td>
<td>37.46/2.67</td>
<td>$y = 0.30x + 226.75$</td>
<td>8.75/18.75</td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>145.9</td>
<td>0.33</td>
<td>-5.44</td>
<td>-3.77</td>
<td>1.29</td>
<td>0.99</td>
<td>35.33</td>
<td>29.05</td>
<td>-</td>
<td>27.50</td>
</tr>
<tr>
<td>$^{d}$POC$_{OC}$</td>
<td>69/22</td>
<td>159.2/128.7</td>
<td>0.34/0.32</td>
<td>-38.03/-60.77</td>
<td>82.38/-58.18</td>
<td>1.17/0.90</td>
<td>0.99/0.84</td>
<td>39.44/33.03</td>
<td>17.11/-9.84</td>
<td>$y = 0.34x + 176.80$</td>
<td>18.75/0.00</td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>152.4</td>
<td>0.33</td>
<td>-43.53</td>
<td>-9.91</td>
<td>1.11</td>
<td>0.97</td>
<td>37.89</td>
<td>10.59</td>
<td>-</td>
<td>18.75</td>
</tr>
<tr>
<td>$^{e}$POC = $a$ * R$_{555}$ + $b$</td>
<td>71/29</td>
<td>158.3/135.6</td>
<td>0.30/0.12</td>
<td>0.00/-52.16</td>
<td>96.80/-23.19</td>
<td>1.37/0.95</td>
<td>1.11/0.96</td>
<td>41.11/36.03</td>
<td>37.07/-5.40</td>
<td>$y = 0.30x + 224.49$</td>
<td>15.00/9.38</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>152.1</td>
<td>0.24</td>
<td>-15.13</td>
<td>21.21</td>
<td>1.25</td>
<td>1.09</td>
<td>39.64</td>
<td>24.75</td>
<td>-</td>
<td>24.38</td>
</tr>
<tr>
<td>$^{f}$POC = $a$ * R$_{667}$ + $b$</td>
<td>71/29</td>
<td>128.2/133.5</td>
<td>0.54/0.27</td>
<td>0.00/7.78</td>
<td>87.67/7.64</td>
<td>1.25/1.08</td>
<td>1.06/1.03</td>
<td>34.26/30.61</td>
<td>25.46/8.06</td>
<td>$y = 0.54x + 147.13$</td>
<td>21.25/15.63</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>129.7</td>
<td>0.46</td>
<td>2.26</td>
<td>11.44</td>
<td>1.20</td>
<td>1.06</td>
<td>33.20</td>
<td>20.42</td>
<td>-</td>
<td>36.88</td>
</tr>
<tr>
<td>$^{g}$POC$<em>{log}$ = $a$ * C$</em>{114}$ + $b$</td>
<td>100</td>
<td>158.6/138.5</td>
<td>0.33/0.27</td>
<td>-36.35/-81.63</td>
<td>91.18/-67.10</td>
<td>1.17/0.82</td>
<td>0.93/0.76</td>
<td>39.57/34.55</td>
<td>16.90/-17.63</td>
<td>$y = 0.35x + 172.72$</td>
<td>3.75/6.25</td>
</tr>
<tr>
<td>$^{h}$POC$<em>{log}$ = $a$ * C$</em>{14}$ + $b$</td>
<td>71/29</td>
<td>140.4/134.9</td>
<td>0.47/0.32</td>
<td>-29.56/-58.78</td>
<td>72.75/-45.53</td>
<td>1.14/0.87</td>
<td>0.90/0.82</td>
<td>36.64/31.54</td>
<td>14.15/-13.48</td>
<td>$y = 0.47x + 139.58$</td>
<td>10.00/9.38</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>138.8</td>
<td>0.42</td>
<td>-38.04</td>
<td>-38.95</td>
<td>1.06</td>
<td>0.86</td>
<td>35.16</td>
<td>6.13</td>
<td>-</td>
<td>19.38</td>
</tr>
</tbody>
</table>

Note. The Rrs667-based algorithms (highlighted in bold) showed the most robust performance in retrieving POC. For each model, the first row represents the statistics of model training/validation, the fitting curve is based on the training data set, and the second row represents the statistics of the whole data set. The last column is the winner percentage of each model based on the POC residuals using pairwise comparison.

$^{a}$POC algorithm based on a linear fit of R$_{bg}$ = Rrs443/Rrs555, $a = -682.1979$, $b = 394.8691$. $^{b}$POC algorithm based on a power fit of R$_{bg}$, $a = 392.1$, $b = -1.096$. $^{c}$POC algorithm based on a linear fit of maximum blue-green band ratio (R$_{bgm}$), only trained for MODIS with R$_{bgm}$ = max(Rrs443, Rrs488/100), $a = -708.9101$, $b = 395.7773$. $^{d}$POC algorithm based on the OCx formula of R$_{bgm}$ (see Equation 2), only trained for MODIS, $a_0 = 4.3219$, $a_1 = -2.7951$, $a_2 = 0.8326$, $a_3 = 0.2756$, $a_4 = -0.1225$. $^{e}$POC algorithm based on a linear fit of R$_{555}$, $a = 78.387$, $b = 4.8837$. $^{f}$POC algorithm based on a linear fit of Rrs667, $a = 593.540$, $b = -66.1469$. $^{g}$POC algorithm based on C$_{114}$, $a = 147.0672$, $b = 2.2594$. $^{h}$POC algorithm based on C$_{14}$, $a = 202.5057$, $b = 2.2377$. 


Figure 7. Same as Figure 6, but performances of different Chlorophyll-a (Chl) algorithms in retrieving Chl using the same training data set (blue dots) and validation data set (green dots), based on the spectral formulas of (a) Rrs555, (b) Rrs667, linear fitting of $R_{bg}$ (linear fitting in (c) and power fitting in (d)), $R_{bgm}$ (linear fitting in (e) and OCx fitting formula in (f)), (g) CI$_{Hu}$, and (h) CI$_{Le}$. The red lines indicate a 1:1 relationship. The statistics shown were based on the whole conjugate data set including both the training data set and the validation data set. The red line is the 1:1 line, the solid blue line is the fitting curve, with ± one standard deviation (blue dashed lines). See Table 4 for more detailed statistics.
<table>
<thead>
<tr>
<th>Regress formula</th>
<th>N</th>
<th>RMSD (mg m⁻²)</th>
<th>R²</th>
<th>MB</th>
<th>Median bias</th>
<th>MR</th>
<th>Median ratio</th>
<th>UPD (%)</th>
<th>MRD (%)</th>
<th>Fitting curve</th>
<th>Winner%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Chl} = a \times \log_{10}(R_{443}) + b )</td>
<td>79/31</td>
<td>3.28/2.34</td>
<td>0.29/0.33</td>
<td>0.00/0.41</td>
<td>2.38/1.01</td>
<td>1.65/1.74</td>
<td>1.24/1.24</td>
<td>61.75/54.92</td>
<td>64.60/74.30</td>
<td>y = 0.29x + 3.67</td>
<td>2.50/9.38</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>3.04</td>
<td>0.30</td>
<td>0.12</td>
<td>0.95</td>
<td>1.67</td>
<td>1.24</td>
<td>59.82</td>
<td>67.33</td>
<td>–</td>
<td>11.88</td>
</tr>
<tr>
<td>( \text{Chl} = a \times (R_{510})^{b} )</td>
<td>79/31</td>
<td>3.23/2.36</td>
<td>0.31/0.30</td>
<td>−0.00/0.20</td>
<td>2.04/1.42</td>
<td>1.66/1.68</td>
<td>1.35/1.26</td>
<td>60.23/55.64</td>
<td>66.34/67.84</td>
<td>y = 0.31x + 3.55</td>
<td>2.50/3.13</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>3.01</td>
<td>0.31</td>
<td>0.05</td>
<td>0.99</td>
<td>1.67</td>
<td>1.30</td>
<td>58.94</td>
<td>66.76</td>
<td>–</td>
<td>5.63</td>
</tr>
<tr>
<td>( \text{Chl} = a \times \log_{10}(R_{510}) + b )</td>
<td>77/21</td>
<td>3.26/2.43</td>
<td>0.30/0.42</td>
<td>−0.00/0.96</td>
<td>2.54/1.67</td>
<td>1.58/2.11</td>
<td>1.25/1.97</td>
<td>60.10/65.26</td>
<td>57.54/111.30</td>
<td>y = 0.30x + 3.64</td>
<td>7.50/3.13</td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>3.10</td>
<td>0.32</td>
<td>0.21</td>
<td>0.95</td>
<td>1.69</td>
<td>1.34</td>
<td>61.21</td>
<td>69.06</td>
<td>–</td>
<td>10.63</td>
</tr>
<tr>
<td>( \text{Chl}_{\text{LC3}} )</td>
<td>77/21</td>
<td>3.35/2.35</td>
<td>0.33/0.36</td>
<td>−0.95/−0.17</td>
<td>2.05/0.70</td>
<td>1.26/1.56</td>
<td>1.03/1.29</td>
<td>54.42/58.66</td>
<td>25.71/56.11</td>
<td>y = 0.32x + 2.63</td>
<td>13.75/25.00</td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>3.16</td>
<td>0.34</td>
<td>−0.78</td>
<td>0.16</td>
<td>1.32</td>
<td>1.11</td>
<td>55.33</td>
<td>32.23</td>
<td>–</td>
<td>38.75</td>
</tr>
<tr>
<td>( \text{Chl} = a \times R_{555} + b )</td>
<td>79/31</td>
<td>3.42/3.30</td>
<td>0.23/0.00</td>
<td>−0.00/−0.02</td>
<td>2.62/−0.63</td>
<td>1.83/1.93</td>
<td>1.09/0.82</td>
<td>64.03/70.28</td>
<td>83.39/92.63</td>
<td>y = 0.23x + 3.98</td>
<td>12.50/12.50</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>3.39</td>
<td>0.15</td>
<td>−0.01</td>
<td>0.25</td>
<td>1.86</td>
<td>1.05</td>
<td>65.79</td>
<td>85.99</td>
<td>–</td>
<td>25</td>
</tr>
<tr>
<td>( \text{Chl} = a \times R_{667} + b )</td>
<td>79/31</td>
<td>2.64/3.35</td>
<td>0.54/0.22</td>
<td>−0.00/1.36</td>
<td>1.73/0.93</td>
<td>1.41/1.87</td>
<td>1.10/1.47</td>
<td>49.43/55.07</td>
<td>41.14/87.13</td>
<td>y = 0.54x + 2.37</td>
<td>25.00/25.00</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>2.86</td>
<td>0.42</td>
<td>0.38</td>
<td>0.52</td>
<td>1.54</td>
<td>1.17</td>
<td>51.02</td>
<td>54.10</td>
<td>–</td>
<td>50</td>
</tr>
<tr>
<td>( \log_{10}(\text{Chl}) = a \times \text{Cl}_{355} + b )</td>
<td>79/31</td>
<td>3.31/3.22</td>
<td>0.33/0.07</td>
<td>−0.93/−1.02</td>
<td>1.79/−1.00</td>
<td>1.27/1.26</td>
<td>1.03/0.76</td>
<td>53.90/65.53</td>
<td>27.18/26.41</td>
<td>y = 0.35x + 2.41</td>
<td>7.50/3.13</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>3.28</td>
<td>0.27</td>
<td>−0.96</td>
<td>0.01</td>
<td>1.27</td>
<td>1.00</td>
<td>57.18</td>
<td>28.97</td>
<td>–</td>
<td>10.63</td>
</tr>
<tr>
<td>( \log_{10}(\text{Chl}) = a \times \text{Cl}_{670} + b )</td>
<td>79/31</td>
<td>2.91/3.77</td>
<td>0.47/0.09</td>
<td>−0.60/−0.36</td>
<td>1.13/−0.32</td>
<td>1.18/1.29</td>
<td>0.97/0.88</td>
<td>43.87/54.71</td>
<td>17.62/28.76</td>
<td>y = 0.54x + 1.76</td>
<td>28.75/15.63</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>3.18</td>
<td>0.34</td>
<td>−0.53</td>
<td>−0.28</td>
<td>1.21</td>
<td>0.95</td>
<td>46.92</td>
<td>20.76</td>
<td>–</td>
<td>44.38</td>
</tr>
</tbody>
</table>

Note: The Rrs667-based algorithms (highlighted in bold) showed the best performance in retrieving Chl.

\( \text{Chl} \) algorithm based on a linear fit of \( R_{555} = \text{Rrs}443/\text{Rrs}555, a = −12.8787, b = 6.4240 \).

\( \text{Chl} \) algorithm based on a power fit of \( R_{555}, a = 6.193, b = −1.258 \).

\( \text{Chl} \) algorithm based on a linear fit of maximum blue-green band ratio \( (R_{670}) \), only trained for MODIS with \( R_{670} = \max(\text{Rrs}443, \text{Rrs}488) \).

\( \text{Chl} \) algorithm based on the OCx formula of \( R_{565} \) (see Equation 2), only trained for MODIS, \( a_{0} = 1.7821, a_{1} = −1.2592, a_{2} = −0.0049, a_{3} = 0.2534, a_{4} = −0.0362 \).

\( \text{Chl} \) algorithm based on a linear fit of \( R_{555}, a = 1379.9, b = −0.3717 \).

\( \text{Chl} \) algorithm based on a linear fit of \( R_{667}, a = 12.352, b = −2.8477 \).

\( \text{Chl} \) algorithm based on \( \text{Cl}_{355}, a = 200.7124, b = 0.3310 \).

\( \text{Chl} \) algorithm based on \( \text{Cl}_{670}, a = 298.7757, b = 0.2852 \).
and 8c), the algal blooms along the coast were associated with Chl of $\sim 3.5 \text{ mg m}^{-3}$ and POC of $300 \text{ mg m}^{-3}$; however, the locally tuned Rrs667-based images of the blooms showed more spatial details and larger concentrations of Chl ($>10 \text{ mg m}^{-3}$) and POC ($>500 \text{ mg m}^{-3}$) (Figures 8b and 8d). Clearly the magnitude of the spring blooms was significantly underestimated using the NASA standard algorithms.

The Rrs667-based Chl and POC algorithms were applied to the MODIS images between 2002 and 2020 and a monthly climatology was derived and compared to the Chl and POC climatology using the default algorithm (Figures 9 and 10). The spatial distribution of the Rrs667-based Chl showed general agreement with the NASA standard Chl, but the magnitude of the Rrs667-based Chl was substantially augmented (Figure 9). In early spring (November), with the reduction in ice cover, phytoplankton Chl started to accumulate on the western shelf. In December and January, when substantial phytoplankton growth and accumulation occur, there is a significant increase of Chl on the western and eastern shelves, and in late summer (February), the Chl remains at a high level near the coast of Victoria Land but decreased in the east. This is consistent with the findings of both individual cruises and climatologies based on in situ observations (Smith & Kaufman, 2018; Smith et al., 2000; 2013).

The monthly climatology of the Rrs667-based POC was also compared to the POC derived from NASA standard processing (Figure 10). Again, the Rrs667-based POC revealed significantly higher values than the standard POC, despite the consistency in spatial distribution between the two POC products. The Rrs667-based POC distributions suggest that in early spring POC increased simultaneously with phytoplankton growth, and by December and January, the areas of high POC regions extended more broadly across the continental shelf and were maximal along the coast. In late summer, high POC concentrations were observed only along the west coast.
Figure 9. The monthly Chlorophyll-a climatology from 2002 to 2020 derived using the NASA standard algorithm (left panels) and the locally tuned Rrs667-based algorithm (right panels).
Figure 10. The monthly particulate organic carbon climatology from 2002 to 2020 derived using the NASA standard algorithm (left panels) and the locally tuned Rrs667-based algorithm (right panels).
4. Discussion

Based on the data collected in the past 20 years, we found that the NASA standard Chl algorithm (Hu et al., 2012) substantially underestimated Chl in the Ross Sea (Figure 3a). One potential explanation for that difference may be that the standard OCI Chl algorithm is trained for the global ocean waters, while our cruise data are restricted to the continental shelf of the Ross Sea, where primary productivity is much higher (Arrigo, van Dijken, & Bushinsky, 2008). However, even in the open waters, we found the NASA standard algorithms also showed an underestimation (Figure S1). One potential reason is that irradiance is a limiting factor for phytoplankton growth in polar regions in spring, and pigment packaging effects could be stronger in higher latitude waters, particularly at higher Chl concentrations (Lewis et al., 2016; Matsuoka et al., 2007, 2011; Mitchell, 1992). Indeed, the specific absorption coefficient (a_{ph}^*(443)) in the Ross Sea were smaller than the global a_{ph}^*(443) (Text S1, Figure S2). The NASA standard Chl algorithm performs well in estimating Chl in the oligotrophic open ocean waters at high latitudes (Haëntjens et al., 2017), but when investigating high biomass regions like the continental shelves, the absolute estimates of phytoplankton biomass may be biased. The strong underestimation might also be caused by the poor representation of the in situ surface measurements for the satellite data over several kilometers in area (i.e., 1 pixel) and a few meters in depth, yet it is a common issue that exists in the matchups between the satellite and in situ data. The direct use of surface measurements in model development should result in more robust algorithms (Lee et al., 2020; Sathyendranath et al., 2019).

We also revisited the NASA standard POC algorithm and other POC algorithms applicable to the Ross Sea with their original model parameterization. These algorithms typically use band ratios (i.e., R_{bg}, R_{bgm}) or band differences (i.e., CI_{Inta}, CI_{Intb}). The severe underestimation of POC using the NASA standard algorithm (Figure 3b) could be attributed to the inapplicability of this algorithm to the region. Indeed, the NASA standard POC algorithm was developed for the global open ocean with POC concentrations ranging between 10 and 270 mg m^{-3} (Stramski et al., 2008). Because of the intrinsic limitations of such empirical algorithms, they can only be applied to oceanic waters with POC levels falling in this range. Yet the POC ranges demonstrated by in situ measurements in the Ross Sea were far above the application range of the NASA standard POC algorithm. The retrieved POC using the standard POC algorithm can differ by 50%-100% from field measurements in coastal oceans (Evers-King et al., 2017). A better approach might be to develop local algorithms for better quantification of the biomass in the Ross Sea. A few studies have tried to locally tune the POC algorithms for waters of the Southern Ocean (Allison et al., 2010; Pabi & Arrigo, 2006). However, based on their original parameterizations, we found that the modeled POC still showed large differences from the measured POC concentrations (see Table 2). To understand this, we analyzed the possible causes for this discrepancy. Specifically, the POC algorithm of Pabi and Arrigo (2006) was developed using Rrs555 based on the relative percentages of P. antarctica and diatoms as determined from observations from two cruises in 1996 and 1997. We repeated this algorithm using our data set based on the qualitative records of which species dominated the water samples. As such, it could introduce some uncertainties in the assessment of this POC algorithm. It also indicates the difficulties of such algorithms due to the temporal and spatial differences in the percentages of each type of phytoplankton. Moreover, based on the multispectral bands from MODIS over a wide range of Chl and POC, we did not find any discriminative spectral characteristics of P. antarctica and diatoms that can be used to differentiate the two (Figure 2). While the size and shape of these groups are markedly different, they also contain very similar pigments that absorb in nearly the same region of the spectrum. The relationship between POC and Rrs555 did not show any distinct difference for either P. antarctica- or diatom-dominated waters. Therefore, for multispectral remote sensing, it may be better not to quantify the POC for a single type of phytoplankton. Like the NASA standard POC algorithm, the POC algorithm of Allison et al. (2010) was developed for the Southern Ocean for surface POC levels ranging from 30 to 120 mg m^{-3}, and that may contribute to the significant underestimation of POC in the productive, high biomass Ross Sea, where POC can reach 1,000 mg m^{-3} (Table 1). Le et al. (2018) tried to improve the global POC algorithm for the coastal waters and proposed a CI-based POC algorithm. Unfortunately, no observational data from the Ross Sea or polar coastal regions were included in their model generalization (see Figure 1 in Le et al., 2018). Therefore, despite the positive correlations between CI_{Int} and POC in the Ross Sea (Figure 4d), we still found a poor retrieval of surface POC from the Ross Sea.
Based on the analyses above, we locally tuned the Chl and POC algorithms using different Rrs formulas. Our results suggest that the Rrs667-based algorithms are by far the most robust in retrieving Chl and POC in the Ross Sea (Figures 6 and 7, Tables 3 and 4). Although the models were tuned based on the observations from the continental shelf of the Ross Sea, they also showed a better performance in the open waters of the Ross Sea than the NASA standard algorithms (Figure S1). Still, one potential disadvantage of the Rrs667-based algorithms is that, they could be sensitive to spectrally dependent errors and noise in satellite-derived Rrs(λ), particularly in coastal and polar regions like the Ross Sea where atmospheric corrections encounter difficulties under high solar zenith angles (He et al., 2018; Wang et al., 2009). In contrast, the CI-based approach is tolerant to the spectral noises to some extent (Hu et al., 2012); however, we demonstrated that the locally tuned CI-based (both Chl and POC) models showed poorer performance. When more field data are available, we suggest that the CI-based approach needs to be re-examined as a means to estimate Chl and POC.

With the improved performance of the Rrs667-based algorithms in estimating Chl and POC, we found that the spring and summer blooms exhibited broader spatial coverages with much more elevated concentrations of particulate matter than those detected in the standard Chl and POC products. This is particularly true in the strong bloom signals in January and February on the western shelf that are captured by the Rrs667 algorithms, but are greatly underestimated using the NASA standard data processing. More importantly, the broader areas associated with significantly higher Chl and POC detected by the Rrs667 algorithms in the surface layer suggest that the productivity of the Ross Sea may be even greater than believed (consistent with the conclusion of Smith & Kaufman, 2018) and thus play an even more important role in the carbon budget of the Southern Ocean.

It has been suggested that as iron-limiting conditions become extreme and widespread in the surface layer of the Ross Sea during austral summer that two things occur: the first being that P. antarctica disappears from the euphotic zone by rapid sinking, and the second being that diatoms continue to grow by reducing the amount of Chl per cell as a response to iron limitation (Ryan-Keogh & Smith, 2021; Smith & Kaufman, 2018). This suggestion appears to be supported by our Rrs667-determined Chl and POC distributions. The magnitude of the POC:Chl ratio temporal changes is substantial (Figure 11). In November this ratio is less than 100, but increases through time, and in February (where substantial spatial variability in both POC and Chl is observed; Figures 9 and 10), it increased significantly (>100). This is actually less than that based on shelf-wide means from discrete observations (Smith & Kaufman, 2018), but suggests that the Rrs667-based determinations of both POC and Chl are substantial improvements over the MODIS-based values. Substantial interannual variations in the ratio also occur, suggesting that iron inputs (either atmospheric or from deeper waters) may vary as well. In contrast, MODIS-based ratios are exceptionally variable, with many values exceeding 200 in spring and early summer, values which have never been measured in situ (Figure 11). While the ratios increase in February, the spatial pattern is the opposite of that found in the in situ climatology. That is, the MODIS ratios are lowest in waters dominated by diatoms (near the coast of Victoria Land) and highest in the central region, whereas observational climatologies show the opposite pattern. The similarity of these ratios between the Rrs667-based estimates and those based on in situ observations strongly suggests that the Rrs667-based POC and Chl values are much more reflective of actual distributions than the previous satellite-based efforts.

5. Conclusions

Our analysis of the current NASA standard Chl and POC algorithms and other algorithms applicable to the Ross Sea suggests that specific algorithms are required for this region to accurately assess the parameters. Using the data collected during 18 cruises over the past 20 years, we found that the Rrs667-based algorithms are the most robust in retrieving Chl and POC in the Ross Sea, and the uncertainties are greatly reduced without obvious biases in the satellite-derived Chl and POC. These estimates would be improved with additional discrete measurements collated with ocean color images. The derived Chl and POC using the refined algorithms showed larger seasonal dynamics than those derived using the current NASA standard processing algorithms. These suggest that the Ross Sea is even more productive than previously estimated and will play an even more important role in regional carbon cycles than thought.
Figure 11. Ratios of POC:Chl generated from the climatologies of each variable in using the NASA standard algorithm (left panels) and the locally tuned Rrs667-based algorithm (right panels).
Acknowledgments
This work was supported by the National Natural Science Foundation of China (NSFC) projects (41906159, 42030708, 41876228 and 41941008). The authors thank many research groups that collected the field data, and the NASA Goddard Space Flight Center (http://oceancolor.gsfc.nasa.gov/) for maintaining and providing the SeaWiFS and MODIS data used in this study. The authors also thank Ocean Data Viewer and Matlab softwares for the data processing and visualization. The authors are grateful for the constructive comments from the two anonymous reviewers that helped improve the manuscript.

References


CHEN ET AL. 22 of 22