remote sensing

MDPI

Article

Achieving Consistent Estimates of Particulate Organic Carbon from Satellites, Ships and Argo Floats

Graham D. Quartly [1,*], Shubha Sathyendranath [1] and Martí Galí [2]

[1] National Centre for Earth Observation, Plymouth Marine Laboratory, Prospect Place, Plymouth, Devon PL1 3JD, UK; ssat@pml.ac.uk
[2] Institut de Ciències del Mar (ICM-CSIC), Pg. Marítim Barceloneta, 37-49, 08003 Barcelona, Catalonia, Spain; mgali@icm.csc.es
* Correspondence: gqu@pml.ac.uk

Highlights

What are the main findings?

- Satellite and in situ POC algorithms agree to within 15%.
- A chlorophyll-based algorithm matches performance of more complex ones.

What are the implications of the main findings?

- Potential for merging satellite and in situ records of POC.
- In open ocean, a chlorophyll-based algorithm will be best, as there are more measurements of this than of optical backscatter.

Abstract

Carbon fluxes from the atmosphere to the ocean and from the ocean surface to the deep ocean are a key pathway in the long-term sequestration of anthropogenic CO_2. Particulate Organic Carbon (POC), which comprises living plankton, detritus and other microscopic organisms, is a very dynamic carbon pool in surface waters, so an ability to assess POC reliably from satellites and autonomous profilers is fundamental to the quantification of the reservoirs and fluxes of carbon within the ocean, and to assess their response to climate change. In situ records from sample filtration during dedicated hydrographic surveys are limited both in terms of spatial coverage and time, so reliable algorithms are required that make use of readily available autonomously collected data that provide much better spatial and temporal coverage. In this paper, algorithms that use ocean colour data from satellites to estimate POC are re-assessed, and then the satellite-derived products are compared with near-surface in situ observations from biogeochemical (BGC) Argo profilers. The satellites and in situ BGC-Argo records match each other to within 30%, but a regional bias persists that may be related to the BGC-Argo fluorometers overestimating the chlorophyll concentration in the Southern Ocean. A simple coarse-resolution regional correction to the observed chlorophyll-a concentration and backscatter coefficient, plus the removal of clear outliers, improves the agreement to approximately 15%. The association of POC with the surface chlorophyll value is so strong that an algorithm based on chlorophyll-a alone provides an almost equally good estimate of POC compared with more complex algorithms that incorporate additional bio-optical variables such as the backscattering coefficient.

Keywords: particulate organic carbon; satellite; ocean colour; BGC-Argo floats; backscatter; chlorophyll; global

Academic Editors: Haidong Pan, Daosheng Wang and Jungang Yang

Received: 27 January 2026
Revised: 4 March 2026
Accepted: 5 March 2026
Published: 9 March 2026

1. Introduction

Sequestration, i.e., the removal of carbon dioxide from the atmosphere to the depths of the ocean, is a complex web of physical, biological and chemical processes [1]. Firstly, there is the two-way transfer of CO_2 between the atmosphere and ocean, mediated by wind, waves, rain, ocean alkalinity, temperature and the partial pressure of CO_2 in the water. Secondly, the dissolved carbon dioxide in the water may be assimilated into organic matter through photosynthesis by phytoplankton, which may, in turn, be consumed by zooplankton and organisms at higher trophic levels. Thirdly, waste excretion, mortality and aggregation lead to a "rain" of organic matter sinking through the water column, which may undergo bacterial respiration and decomposition or some fraction of it may reach the bottom intact. There is also vertical transport through the mixing and physical advection of water masses. Phytoplankton is a variable component of the Particulate Organic Carbon (POC) found in the pelagic ocean, with detrital material and other living organisms making up the rest of the POC pool.

Chlorophyll-a (Chl-a), the main pigment contained in phytoplankton, is a commonly used measure of phytoplankton concentration. The direct measurement of chlorophyll is a time-consuming process involving the collection of water samples and their filtration, followed by the assay itself, leading to large gaps in the data coverage. The remote sensing of ocean colour using satellites, which use algorithms to infer Chl-a from reflectances measured at multiple wavelengths in the visible domain, is therefore an important tool for extrapolating the sparse in situ observations to large scales. Ocean colour methods exploit the absorption characteristics of Chl-a to underpin the satellite-based algorithms. The fluorescence response of phytoplankton to broad-band excitation by visible light from the sun, or to single wavelength excitation on autonomous observing devices, also serves as an indicator of Chl-a concentration. However, various factors, such as observation conditions, the physiological status of the phytoplankton and the species composition of the plankton community, can influence the conversion between fluorescence and chlorophyll concentration. The intercomparison of absorption-based and fluorescence-based tools for monitoring Chl-a is also not trivial, because algal physiology has a stronger influence on the fluorescence signal [2] than on phytoplankton absorption coefficients [3–5].

Particulate Organic Carbon can also be estimated from satellite-observed radiances, with a number of algorithms being proposed (see a review of various types of algorithms in Evers-King et al. [6]). Satellite-based algorithms can only provide measurements within the first optical depth, within which the light level drops by a factor of e^{-1} compared with its surface value. For the wavelengths of interest, this depth typically ranges from ~5 to ~40 m depending upon water clarity. Consequently, complementary tools are needed to observe POC at depth; in fact, recent papers [7,8] have espoused estimating POC from optical backscatter at 700 nm, $b_{bp}(700)$, measured by biogeochemical (BGC) Argo floats.

The core Argo programme co-ordinated the deployment and processing of autonomous profilers that provide physical measurements of conductivity, temperature and depth every 10 days from the surface down to 2000 m. The original objective of near-global uniform coverage of physical measurements with 3000+ floats has been continuously met since November 2003. Recognising the potential offered by such an autonomous platform, effort has been spent on developing a host of biogeochemical sensors that can augment the physical sensors on these floats. There are presently a few hundred such floats in operation (see Figure 1), although coverage is far from uniform as they have been deployed with various scientific goals in mind rather than with the aim of establishing a quasi-global homogeneous operational dataset. The synergy between ocean colour and BGC-Argo floats is well-recognised, and further improvement is identified as a priority [9].

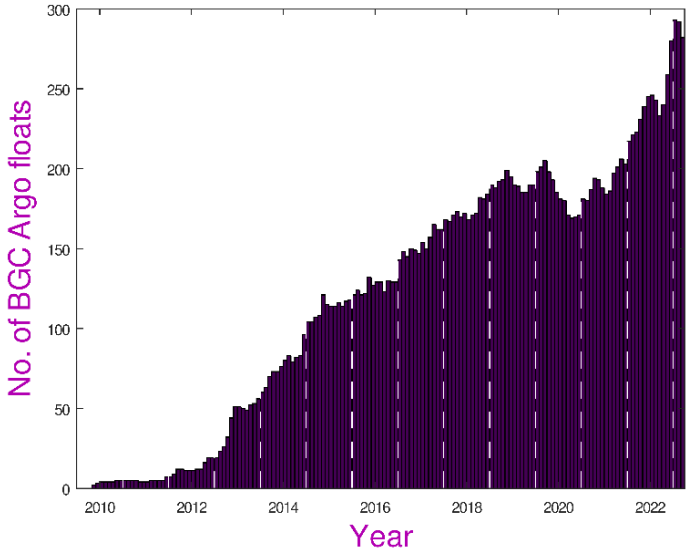

Figure 1. Number of active BGC-Argo floats providing profiles within each month.

This paper seeks to bring together estimates of POC from satellites, BGC-Argo floats and shipborne observations to produce a consistent dataset spanning more than a decade. This could then be exploited to understand the drivers of fluxes and allow an initial look at interannual changes, whether in response to climate modes such as ENSO or long-term climate change. To develop this dataset, we need to construct algorithms for POC from satellite and Argo data that are unbiased with respect to the limited in situ observations, and consistent with each other. In Section 2, we introduce the various sources of data, and their coverage, with Section 3 providing the intercalibration between them, culminating in POC algorithms for both satellite and Argo float data that are consistent with the in situ measurements. Section 4 then discusses the slight regional biases that exist between algorithms, with Section 5 providing a summary of the findings.

2. Data and Methods

There are three types of data that are used together in this study—those from ship, satellites and BGC-Argo floats. The ship data on POC [10] are a collation of near-surface observations from a multitude of oceanographic research cruises. The data coverage is quasi-global and spans many decades, with the POC measured using a broadly consistent protocol involving the filtration of near-surface water samples, drying and the removal of the inorganic component using acid (see Section 2.1 of Evers-King et al. [6]). Kong et al. [10] had an initial dataset of 5972 observations; after the removal of inland records and those for which no synoptic satellite ocean colour data were available, there remained 3288 records spanning 1997 to 2020, with a large proportion of the observations coming from the Atlantic Meridional Transect programme (AMT, see Figure 2).

The first ocean colour sensor was the Coastal Zone Colour Scanner, launched in 1978, but continuous global records of ocean colour began with the launch of SeaWiFS in 1997, with several different sensors being launched over the ensuing decades, with generally increasing performance in terms of spatial resolution, swath width, number of bands and radiometric precision. The European Space Agency's Ocean Colour Climate Change Initiative (OC-CCI) programme has supported work to homogenise data from these different instruments to create a self-consistent multi-decadal dataset of optical water measurements, including remote sensing reflectances and various derived bio-optical products, including Chl-a and particle backscattering coefficients and derived products. We used the latest version of the OC-CCI data (v6.0), which, at the time of access, spanned September 1997

to April 2023. The data used here were daily files with a 0.042° × 0.042° grid that use all available ocean colour sensors to provide an average of the reflectances at multiple wavebands (Figure 3), along with the derived products; see Sathyendranath et al. [11] for more details.

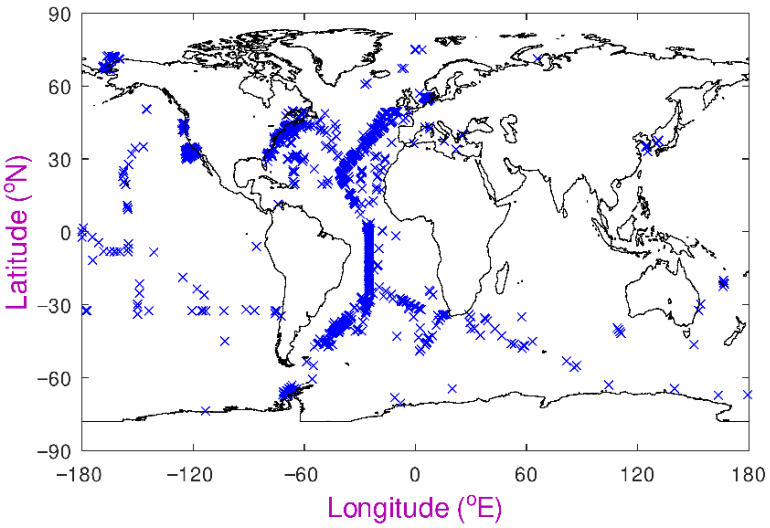

Figure 2. Locations of the near-surface in situ records of POC in the Kong et al. database [10].

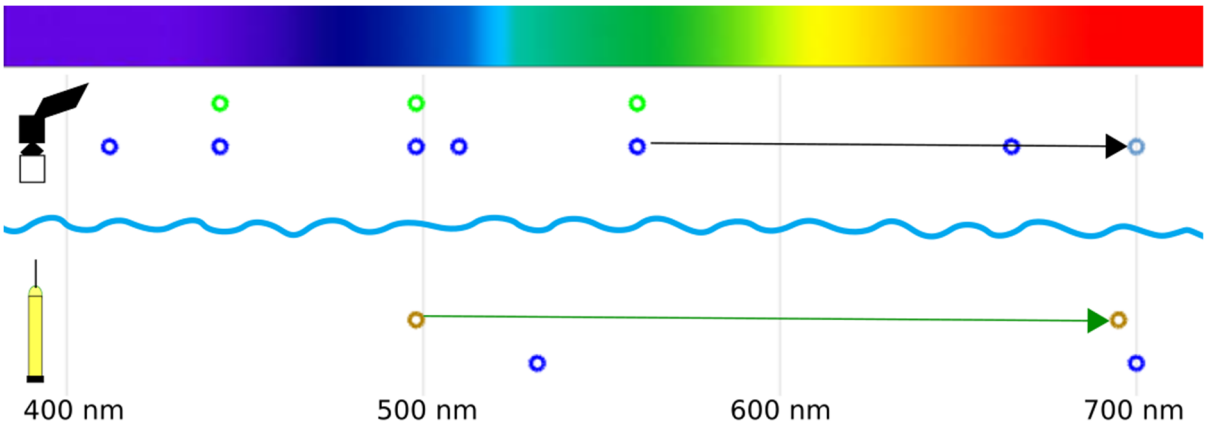

Figure 3. Illustration of the different wavelengths used in remote sensing by both ocean colour satellites and BGC-Argo profiling floats. The Ocean Colour CCI project processes ocean colour data from multiple sensors to retrieve the sea surface reflectance (R_{rs}) at the various wavelengths that past and present visible light sensors have used. From these are calculated the particulate backscatter coefficient (b_{bp}), and absorption coefficients for absorption by dissolved and detrital (a_{dg}), phytoplankton (a_{ph}) and total (a_{tot}), at six key wavelengths (412, 443, 490, 510, 560 and 665 nm, indicated by the blue circles). Of these, those also marked in green are used to produce an estimate of chlorophyll concentration (Chl-a). The diffuse attenuation coefficient at 490 nm, k_d490, is also calculated from R_{rs}. In this work we also use the backscatter products at 560 and 665 nm to extrapolate to the value at 700 nm. For sub-surface remote sensing from BGC-Argo floats, chlorophyll is determined from in vivo fluorescence, using an excitation pulse at 470 nm, with the resultant emission being recorded at 695 nm. BGC-Argo floats also usually record the optical backscatter at 700 nm, with a few profilers also measuring it at 532 nm (blue circles).

The BGC-Argo data were downloaded from the Coriolis data server (https://www.coriolis.eu.org/Observing-the-Ocean/ARGO (accessed on 23 March 2023)). The data fields used were BBP700 and CHLA_ADJUSTED, where the latter name indicates some delayed mode processing to bring values more in line with contemporary shipborne measurements. These are both optical measurements, but at different wavelengths from

the satellite observations (Figure 3), and based on different properties, i.e., backscatter and fluorescence, rather than radiances. In this investigation the quality flags for b_{bp} were ignored, in order to obtain the greatest amount of data, with the data subsequently assessed for anomalies and outliers disagreeing with other observations. The profile depths are not evenly spaced throughout the water column; therefore, to ensure a good representation of the vertical structure, profiles were only kept if there were at least 50 measurements with the deepest being at least 150 m below the surface. (This enabled us to check that each profile had near-zero values at depth, and that there were no strong biases in the sensors.) A simple five-point median filter was then applied to remove outliers, since large values of $b_{bp}(700)$ may exist due to the occasional presence of aggregates of particles or zooplankton. Finally, the data were linearly interpolated to a regular 1 m vertical grid. In many cases there are no valid measurements in the top few metres, so a "surface" value was defined as the average of the interpolated values at 5, 6 and 7 m. As many of the BGC-Argo floats have no data in the top few metres, this definition of a "surface value" is a compromise aiming to maximise the number of useful comparison points that coincide with the optical depth observed by the satellites. Clearly the existence of sharp stratification could lead to some discrepancy between the layers observed by satellite and profiler. [Notably, in this work we only use these "surface" measurements from the BGC-Argo floats, but the selection of profiles with records below 150 m helps ensure that the optical sensors are performing as expected].

3. Results from Intercomparing Datasets

In this paper, we compare a number of different datasets and algorithms against each other. Our analysis focuses on the ability to observe consistent features in different datasets, i.e., whether they show the same spatial or temporal patterns, rather than on the absolute magnitude of the estimated POC. Therefore, in most of the analyses we consider $\log_{10}(\text{POC})$, so that the fractional changes at low POC values are considered as important as at high POC values, and we also make a linear adjustment to these logarithms, so that the best-fit line is $y = x$. This means that comparisons are made on the degree of correlation (r^2) and the root mean square difference (RMSD), without that being affected by the gross underestimation or overestimation of a particular dataset or algorithm. In the end, we bring all these adjustments or "tuning" together to result in algorithms that are unbiased with respect to the in situ data.

Ideally, one would have a fully calibrated POC dataset based on BGC-Argo measurements that were also compatible with shipborne and satellite-based estimates. As there were only a few direct matchups of BGC-Argo profiles with ship data, our approach was to first assess satellite records through comparison with near-surface in situ data and then compare BGC-Argo data with satellite estimates.

The range of POC values encountered in the ocean spans more than two orders of magnitude. In order to develop algorithms that would reveal spatial and temporal variations at both low and high concentrations, we worked exclusively with logarithms of the values from the ships, satellites and BGC-Argo floats, rather than using absolute values, which would lead to algorithms that matched the high values particularly well. Fitted lines are calculated using "Model 2" statistics, i.e. assuming that there are equal errors in both sets of measurements, rather than taking one set as the "ground truth".

3.1. Assessment of Satellite Remote Sensing Algorithms

There are several published algorithms for inferring POC from R_{rs} or derived products. Here we considered those proposed by Loisel et al. [12] and Stramski et al. [13]:

$$\text{Loisel: POC} = (400/0.0096)\, b_{bp}(490)\, \text{Chl}^{0.253} \quad (1)$$

$$\text{Stramski I: POC} = 203.2\, (R_{rs}(443)/R_{rs}(555))^{-1.034} \quad (2)$$

$$\text{Stramski II: POC} = 308.3\, (R_{rs}(490)/R_{rs}(555))^{-1.639} \quad (3)$$

For comparison we used a matchup database compiled by Kong et al. [10] of ship-based near-surface POC values, with the OC-CCI remote sensing radiance values updated from v4.2 to v6. These were assessed by fitting lines to log–log plots, identifying and removing outliers and fitting a line to the remaining data. The number of outliers was broadly similar (see column 2 of Table 1), reflecting the fact that these were probably mainly due to issues with the ground truth data (a few extreme values were in Chesapeake Bay, and a number of others came from a single AMT cruise, for which procedures may have been different). On the basis of the r^2 value and the fewest outliers, the Stramski II algorithm was deemed the best. However, 555 nm is not one of the standard wavelengths in the OC-CCI dataset, so we used the 560 nm channel and tuned the coefficients appropriately:

$$\text{(modified): POC} = 320\, (R_{rs}(490)/R_{rs}(560))^{-1.579} \quad (4)$$

Table 1. Fitting statistics for various satellite ocean colour algorithms assessed against shipborne measurements of POC. There were 3288 matchups.

Algorithm	No. of Outliers	Slopes	r^2	RMSD
Loisel	224	0.928	0.938	0.116
Stramski I	232	0.933	0.957	0.097
Stramski II	223	1.006	0.963	0.100
modified	224	1.000 [1]	0.963	0.099

[1] Note that, as the modified Stramski was developed with these data, the slope is 1.0 by design.

This led to similar results to Stramski II (see Table 1), and is thus implemented for the satellite data throughout the rest of the paper.

Figure 4 illustrates the benefits of working exclusively in log space, as the scatter about the best-fit line appears roughly similar across the dynamic range, whereas in absolute values, the typical size of the deviations from the line is greater at large values (Figure 4a), i.e., it is heteroscedastic. Whilst the switch to logarithms does result in an increased r^2 value, it also helps highlight some deficiencies in the data. First, the satellite-derived $\log_{10}$POC values do not go below 1.34 (POC = 22 mg m^{-3}), indicating that very low POC values are a challenge for satellite-based algorithms. Therefore, de facto, we accept that concentrations below 22 mg m^{-3} are undetectable using these satellite algorithms. Second, there is a cluster of outliers lying below the main distribution. Some of these are in extremely coastal waters, which tend to be regions of great spatial heterogeneity. Our assessment procedure, involving identifying and discarding outliers more than three standard deviations from the line, removes 224 such points (representing 6.8% of all the observations), and leads to the reduction in RMSD from 0.150 to 0.099 and the increase in r^2 from 0.918 to 0.963 (which are the values given in Table 1).

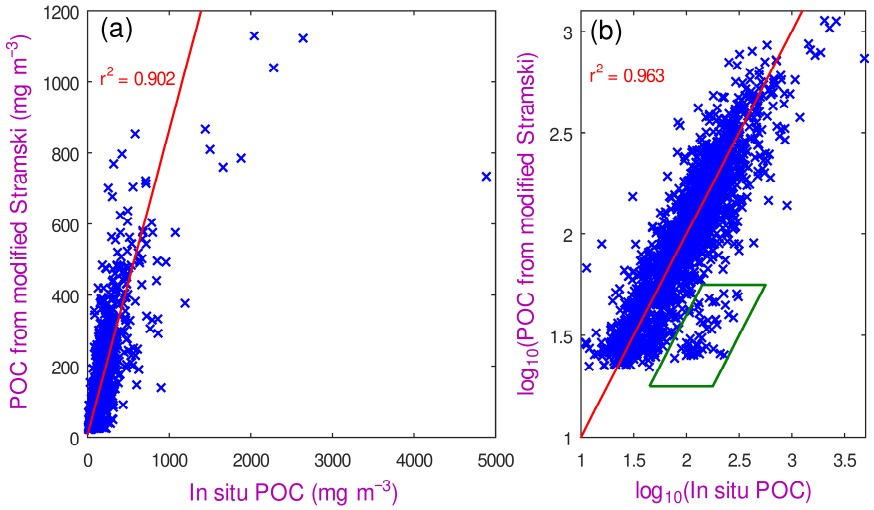

Figure 4. Scatter plots of 3288 in situ POC values against estimates from the modified Stramski algorithm, displaying in (**a**) absolute values and (**b**) logarithms. Some of the extreme in situ values were recorded in Chesapeake Bay; another cluster of outliers (in the green rhombus) are for satellite estimates of $\log_{10}$POC more than 0.3 below the in situ values.

3.2. Matching Satellite and BGC-Argo Data to Assess Relative Performance

We explored two independent algorithms that seek to relate BGC-Argo optical signals to the POC concentration. The first is based on b_{bp} alone, following the approaches first proposed by Loisel et al. [14] and Cetinić et al. [15], and will hereafter be termed "LC". Loisel et al. [14] developed an algorithm that was linear in b_{bp} at 555 nm, and Cetinić et al. [15] expanded the idea using a much larger dataset of b_{bp} observations at 700 nm, which is relevant for the BGC-Argo floats. They found negligible differences between diatom and non-diatom communities, but noted differences between sampling strategies, and between the mixed layer and below it. In a posterior analysis, Galí et al. [7] found very similar POC vs. $b_{bp}(700)$ linear slopes in surface waters of the Mediterranean and subpolar North Atlantic, whereas smaller and larger slopes were found, respectively, in the Southern Ocean and in subtropical gyres. In this work, we adopt the coefficient used by Galí et al. [7] for the subpolar North Atlantic (see Equation (5)), and apply it for surface estimations in all regions. (As our subsequent analysis involves rescaling estimates from algorithms to match observations, the initial scale factor is unimportant; the key thing is that the LC estimate is a monotonically increasing function of $b_{bp}(700)$ alone.)

$$\text{LC: POC} = 41{,}550\ b_{bp}(700) \tag{5}$$

where this algorithm is only being applied to surface values of b_{bp}. The second algorithm investigated is that of Koestner et al. (2022) [8], hereafter "K22", which is a function of both b_{bp} and Chl-a, with the latter acting as a constraint on how large the POC value could be. The K22 algorithm is given by

$$\text{K22: POC} = 89.423\ b_{bp}(700)^{0.1881}\ (\text{Chl-a}/b_{bp})^{\gamma} \tag{6}$$

$$\text{If POC} < 33.4\ \text{POC} \rightarrow 1.636\ \text{POC} - 21.2 \tag{7}$$

where $\gamma = 0.7591 + 0.1934\ \log_{10}(b_{bp})$. The magnitudes produced by the LC and K22 algorithms differed by more than a factor of two. To explore the effect of the choice between these two on the distribution of the outputs rather than just on the magnitude, the

logarithm of the output from the LC was rescaled to match the range of the logarithm of the K22 estimates:

$$\text{rescaled LC: } \log_{10}(\text{POC}_{\text{mod\_LC}}) = 0.7781\ \log_{10}(\text{POC}_{\text{LC}}) + 0.5472 \quad (8)$$

which is equivalent to making the original algorithm proportional to b_{bp} ^ 0.7781. Figure 5 shows a comparison of this rescaled LC with the K22 algorithm for all the locations for which BGC-Argo observations are available. POC estimates from the scaled LC increase monotonically with b_{bp}, whereas for the K22 algorithm low Chl-a values limit its estimates of POC, such that for large b_{bp} values an associated very low Chl-a value will reduce the inferred POC estimate by as much as a factor of ten (one unit in the logarithm scale). Although the LC algorithm has been rescaled so that its best-fit line with respect to the K22 estimates is the 1:1 line, there are strong differences at the extremes, with the rescaled LC producing higher peak values, whilst the K22 algorithm returns much lower POC estimates than the rescaled LC at very low b_{bp} values. We proceed by using a rescaled LC algorithm alongside the K22 one in order to investigate how the additional information from the Chl-a values affects the geographical patterns and the comparisons with in situ and satellite data.

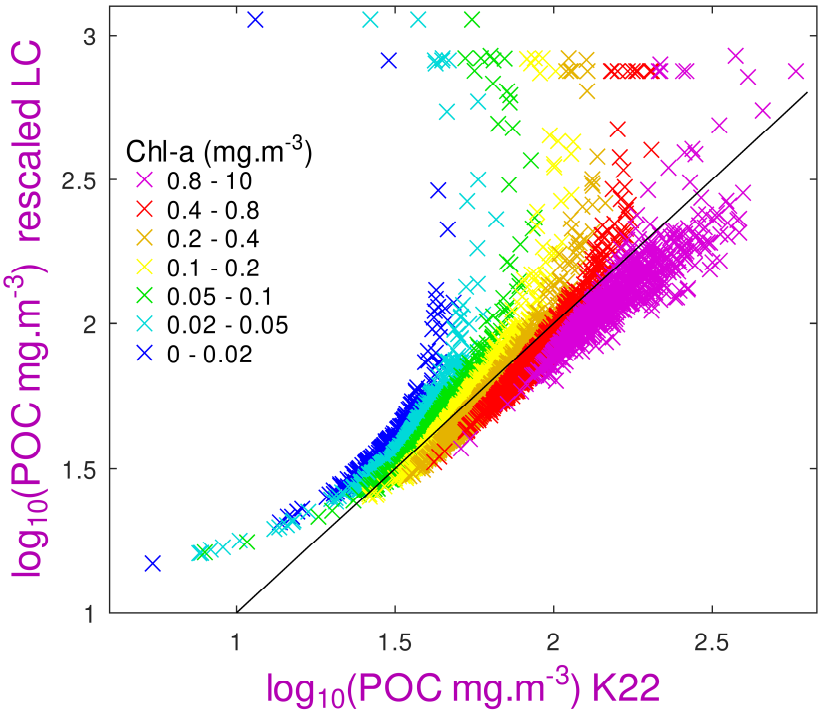

Figure 5. Scatter plot of surface POC estimates from K22 and the rescaled LC algorithms), segregated according to chlorophyll concentration (Chl-a), ranging from dark blue for the lowest Chl-a to pink for the highest values. The black line shows the 1:1 relationship, which is the simple "Model 2" linear fit to all the data (because that was the intention of the rescaling in Equation (8)).

Next, we compare the output of both algorithms with values derived from the satellite data. For this, we match the BGC-Argo surface observations with ocean colour data from the OC-CCI project, and apply the modified Stramski algorithm (Equation (4)) to them. To enable a large number of matched points for validation and analysis, the proximity criteria were relaxed from the exact pixel and same day to a median of observations within ±2 days and ±1 pixel. Appendix A shows that this greatly increased the number of matchups for comparison, whilst only adding a little uncertainty to the satellite values.

Figure 6 shows the correspondence of POC estimates from satellites with those from the rescaled LC and K22 algorithms applied to BGC-Argo surface data. There are

36,886 matchups with valid estimates from the Stramski, LC and K22 algorithms. In this analysis, the two BGC-Argo-based algorithms are linearly adjusted in log space so that the best-fit lines with respect to the satellite estimates are the $y = x$ lines. This does not imply that the satellite estimates are "more correct" but simply allows us to focus on the regional and temporal variations without being distracted by differences in scaling. The scatter plots in Figure 6a,c depict one sixth of the points, whilst the population density plots (Figure 6b,d) use all the points and show how they are concentrated close to the 1:1 line.

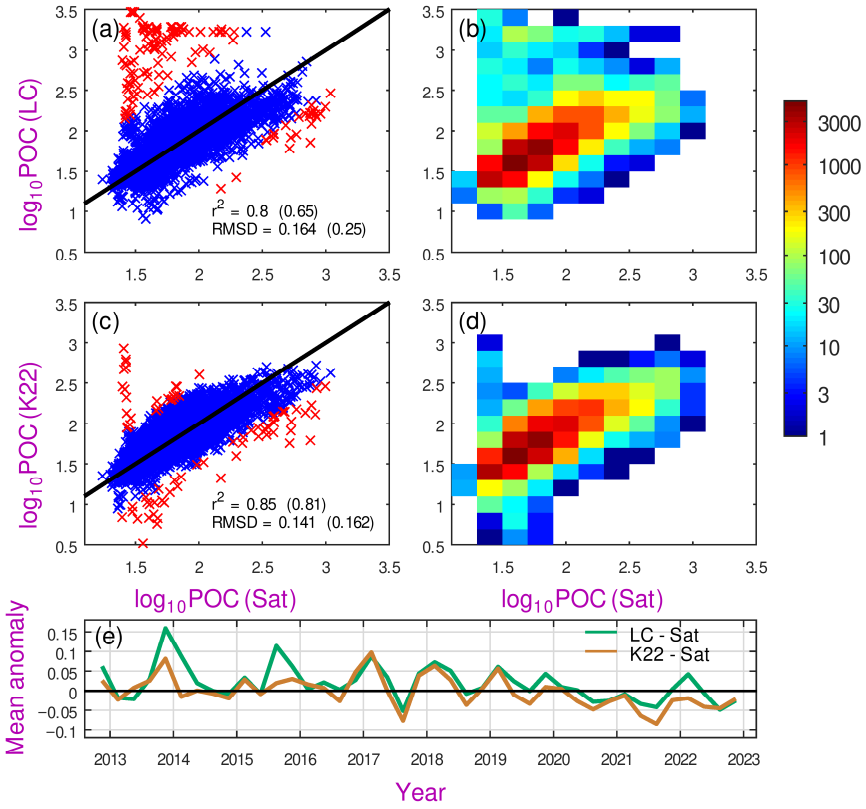

Figure 6. Comparison of the two b_{bp}-based algorithms with matched satellite data. (**a**) Scatter plot and (**b**) population density plot showing correlation of LC with modified Stramski applied to satellite data. The crosses only show one sixth of all the points, with the red ones indicating those that lie more than three S.D. from the best-fit line and the blue crosses being the "good" matchups. The first value for r^2 and RMSD are for the good data, with those in parentheses being the initial results when all data were used. Algorithms have been adjusted so that best-fit line is the 1:1 line. (**c**,**d**) are the corresponding plots for the K22 algorithm. (**e**) The mean bias (BGC-Argo estimates from the LC and K22 algorithms minus satellite estimate from the modified Stramski algorithm) calculated for each 3-month period.

The points in red (in the left-hand panels of Figure 6) indicate those more than three standard deviations (S.D.) away from the mean line, and are classified as outliers. A number of these outliers are noted for low satellite values (~1.5 in $\log_{10}$ units) but with $\log_{10}$BGC-Argo values above two. Roughly half of those corresponded to just two Argo floats in the middle of the tropical Atlantic, and could potentially be due to water ingress affecting a long series of profiles. (An instrument malfunction is deemed more likely than episodic atmospheric deposition as the anomalous readings persisted over many months. In future work the quality flags on the BGC-Argo data will be heeded to help address whether anomalous values can be definitely attributed to problems with the optical sensor.) For the LC algorithm there are also some additional outliers for $\log_{10}$POC estimates around 3.2; these are not present for the K22 algorithm, and are likely to be in low Chl-a waters, where that acts as a constraint on K22's estimates (see Figure 5). The automatic outlier

detection flags ~1.9% of points in Figure 6a, and ignoring them reduces the root mean square difference (RMSD) from 0.250 to 0.164, with an increase in r^2 to 0.80. Only 1.3% of points are flagged in Figure 6c for the K22 algorithm, with the RMSD being reduced from 0.162 to 0.141 and r^2 improved to 0.85.

We now consider whether any particular locations or periods were associated with biases between these algorithms. Figure 6e shows a time series of the mean anomaly (BGC-Argo estimate minus satellite value) for each 3-month period. Although there may seem to be a slight trend, this is not significant given that the spatial sampling is far from uniform and varies with time according to various deployment campaigns. Figure 7 explores the bias between BGC-Argo algorithms and satellite estimates in a spatiotemporal context. Broadly, there is a latitudinal variation in the bias, with both the BGC-Argo algorithms yielding lower values than the modified Stramski (satellite) algorithm in the northern high latitudes and reading higher in the Southern Ocean (south of 30°S). Notably, the implementation of LC by Galí et al. [7] used lower coefficients for the Southern Ocean.

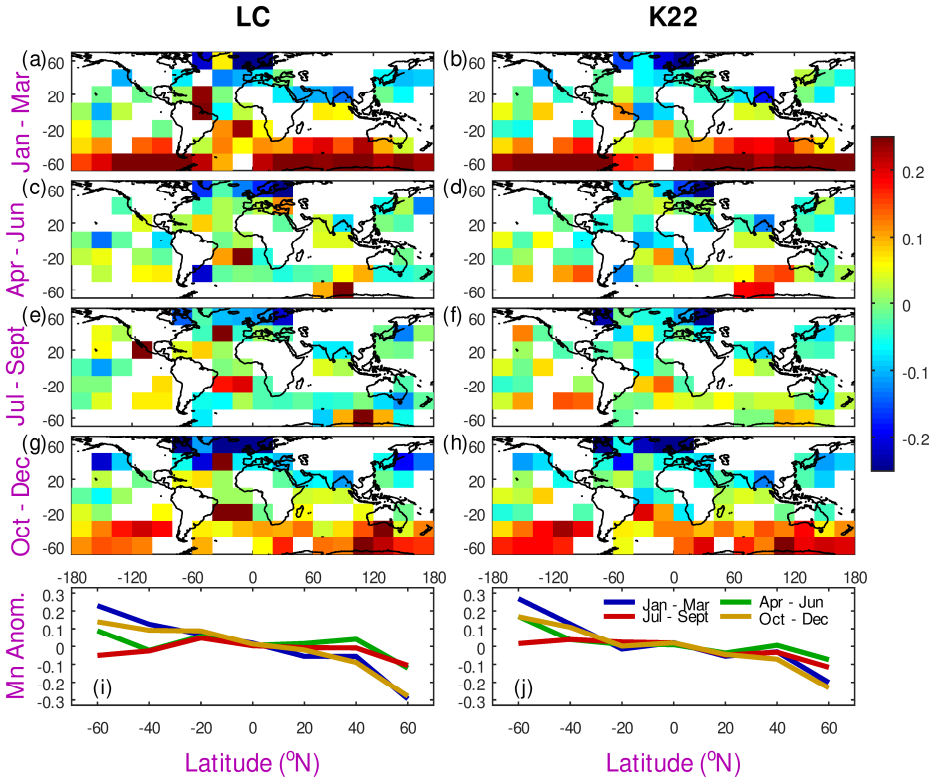

Figure 7. Seasonal variation in the bias (the mean of BGC-Argo minus satellite) according to the two b_{bp} algorithms. The lower plots show the latitudinal average, clearly depicting the consistent pattern of underestimation in high northern latitudes and overestimation in the Southern Ocean. (**a**) Mapped bias of LC algorithm for January-March, (**c**) for April-June, (**e**) for July-September, (**g**) for October-December. Panel (**i**) brings together the latitudinal variations for these 4 seasons. Panels (**b**,**d**,**f**,**h**,**j**) show the corresponding plots for the K22 algorithm.

For the latter region the effect is most pronounced in January–March and then October–December, i.e., over the austral summer season. In a similar vein, the boreal summer in the subarctic Atlantic has a higher bias (less negative) than the other half of the year. This could be due to low viewing angles affecting the satellite ocean colour values in winter or a property of the summer phytoplankton community structure in haline-stratified waters. Gali et al. [16] has suggested that deep convection processes in the high-latitude oceans may deplete POC reserves to levels too low for accurate retrieval from satellite data. As the standard satellite retrieval does not go below $\log_{10}$POC = 1.35 (see Figure 4b),

this could induce a negative bias (BGC-Argo estimate minus satellite) for wintertime. However further work is required to fully understand the seasonal variation in bias at these high latitudes. Satellite records of coccolithophore blooms [17] were examined to see whether they were affecting the satellite estimates in summer, but no connection was found. These latitudinal variations prompted the re-examination of the data used as input for the algorithms, specifically a known overestimation of Chl-a concentration by BGC-Argo floats [2]. In Appendix B we examine both the $b_{bp}(700)$ and Chl-a estimates, contrasting them with the matched satellite data, with the satellite b_{bp} at 700 nm being determined from an extrapolation of the values at 560 and 665 nm wavelengths.

3.3. Correction for Regional Chlorophyll Overestimation

A pointwise comparison of POC estimates from the LC algorithm against the satellite records (Figure 8a) yields an r^2 value of 0.824 after the outliers are removed. However, there are clearly a number of high values for LC ($\log_{10}$POC > 2.4) that do not match well with the satellite estimates. A higher r^2 value is obtained for the K22 algorithm, as its use of Chl-a values has significantly reduced those high estimates. In Appendix B we note that the BGC-Argo $\log_{10}(b_{bp})$ values are lower than the corresponding satellite estimates by about 0.2, except for at very low values, with the bias varying little with Chl-a or the region. On the other hand, the difference between the BGC-Argo and satellite estimates of Chl-a does appear to be a function of Chl-a, b_{bp} and the region. Figure 8c,d show the effect of adjusting for these apparent biases and then reapplying the LC and K22 algorithms. For the LC algorithm, the applied change in b_{bp} just reduces the range of the estimates, without improving r^2. Whilst the b_{bp} change also reduces the range of outputs for the K22 algorithm, the changes in the input Chl-a values, especially with region, reduce the scatter and thus improve r^2.

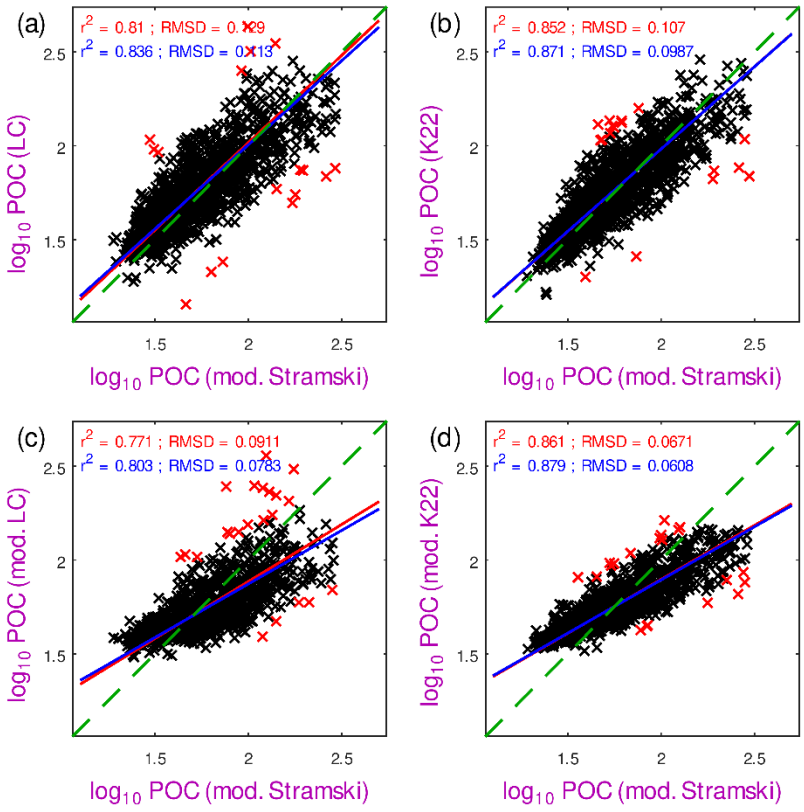

Figure 8. *Cont.*

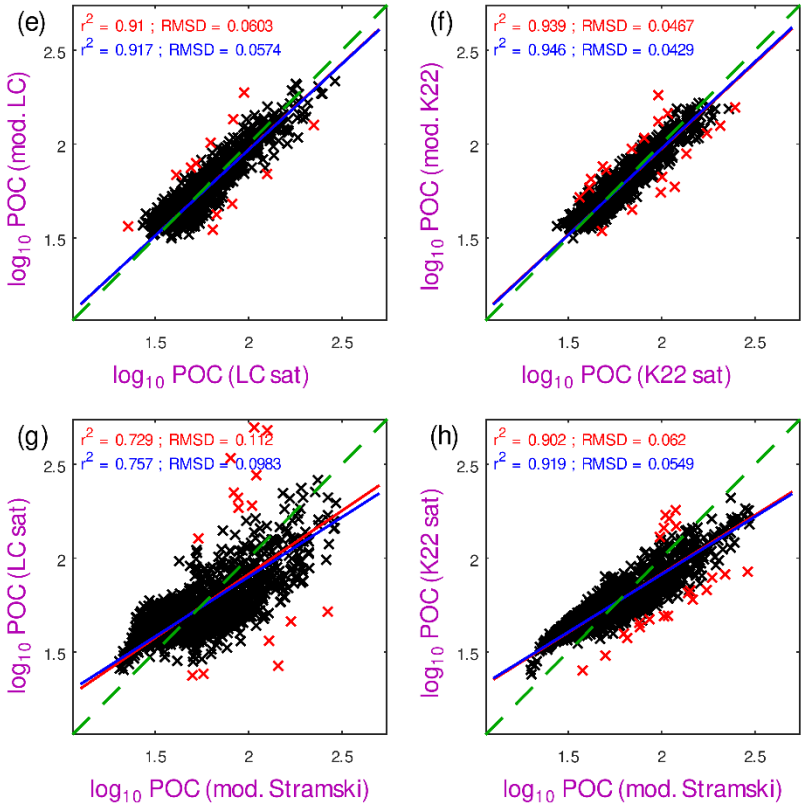

Figure 8. Comparison of POC from various satellite and BGC-Argo algorithms, with only every tenth point plotted for clarity. The left-hand panels use the LC algorithm and the right-hand one uses K22. The fitted straight lines were calculated using Model 2 statistics, i.e., considering error equally in y- and x-terms, with the red lines using all the data, and the blue lines being the fits once the outliers, depicted by red crosses (typically 2%), were excluded. (**a**,**b**) compare the two BGC-Argo algorithms with satellite estimates using the modified Stramski algorithm (Equation (4)). The RMSD values in (**a**) are 0.134 and 0.122. (**c**,**d**) show the results when the Argo algorithms are applied to adjusted values of b_{bp} and Chl-a (see Appendix B). (**e**,**f**) compare BGC algorithms with the same algorithms applied to satellite estimates of b_{bp} and Chl-a. (**g**,**h**) show the variation due solely to the algorithms by comparing BGC algorithms applied to satellite data with the modified Stramski applied to satellite data. In all cases the dashed green line represents $y = x$.

The third row of Figure 8 contrasts the results for the same BGC-Argo algorithm applied to the satellite and float data, thereby showing the effect of discrepancies in the estimates of b_{bp} and Chl-a. This shows that if consistent algorithms are applied to satellite and float data that have been homogenised, the difference in the derived POC values will be less than 0.06,, i.e., corresponding to absolute values of less than 15%. The last row shows the analysis when different estimates are made from the satellite data i.e., applying the LC or K22 algorithm to satellite estimates of b_{bp} and Chl-a and comparing the standard satellite algorithm (Equation (4)) that is based on different wavelengths to that used for the b_{bp} and Chl-a estimates (see Figure 3). This shows that there is significant scatter between these synoptic estimates due to the assumptions behind each of these POC algorithms.

Figure 9 shows how the algorithm comparisons in Figure 8 vary with latitude. By comparing the green lines with the blue ones, we note that the strong bias noted in the Southern Ocean is greatly reduced once the BGC-Argo b_{bp} and Chl-a values have been adjusted to match satellite-derived values. The yellow lines indicate that the Argo-based and satellite-based algorithms produce very consistent estimates when both are applied to the OC-CCI radiances, with minimal latitudinal variation. Interestingly the red lines

in Figure 9a do show a slight regional bias between these algorithms both applied to satellite data.

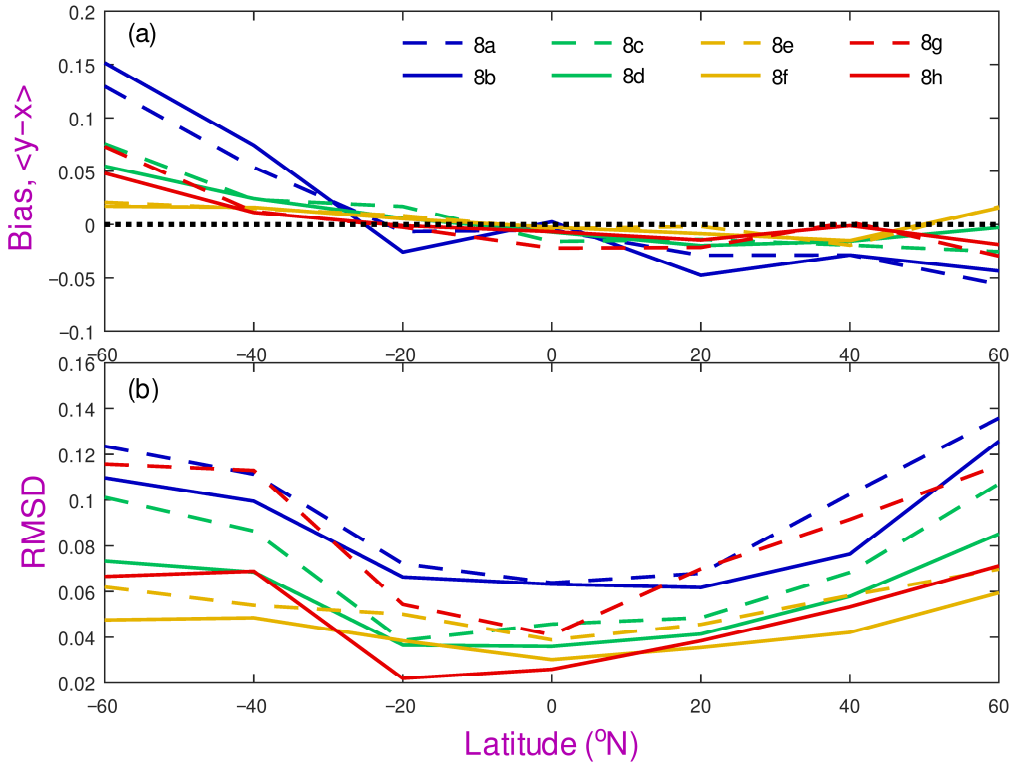

Figure 9. Portrayal of the various POC comparisons in each of the panels in Figure 8 as a function of latitude. Anomalies in $\log_{10}$POC are calculated as the satellite-derived measure (y) minus the Argo-derived estimate (x), using Model 1 statistics (i.e., using fits partitioning all the error in the y-component, with (**a**) showing the mean bias in each 20° latitudinal band, and (**b**) showing the root mean square difference (RMSD). All the dashed curves represent comparisons using the LC algorithm, and the solid ones represent the K22 algorithm. [Note that the biases for the LC and K22 algorithms are very similar in all comparisons, indicating that differences between the BGC-Argo and satellite retrievals of Chl-a and $b_{bp}(700)$ dominate uncertainty in POC estimates.].

Finally, we complete the set of intercomparisons by contrasting these BGC-Argo-based algorithms with the in situ observations of POC. Note again that, as close matchups of BGC-Argo with in situ measurements of POC are rare, this comparison is performed using the LC and the K22 algorithms applied to the satellite observations of $b_{bp}(700)$ and Chl-a at the locations of the in situ observations. Other satellite-based algorithms are also assessed in an identical manner to provide contrast. Figure 10 shows six comparisons of satellite-based estimates with the in situ dataset. In this case there are 3281 points spanning a full range of latitudes, chlorophyll and POC concentrations. Each comparison identified a number of outliers that lay significantly far from the main envelope; 82 of these points were common to three or more of the algorithms, and are plotted in grey but not used in the calculation of the summary statistics. The majority of these outliers correspond to relatively high in situ values, and may indicate errors in the "ground truth" or in the satellite data, or they may represent some oceanic conditions with high spatial heterogeneity.

All the satellite algorithms have been linearly adjusted so that the best-fit line (ascribing errors equally to in situ and satellite estimates) is the y = x line shown in red. Therefore, the useful measures are the r^2 value, the RMSD, and whether the differences between those two estimates are spread uniformly or whether they are more pronounced in certain conditions. The first panel (Figure 10a) shows the results for the modified Stramski algorithm, based on satellite radiances, and provides the best r^2 and RMSD values of all

six algorithms shown. The LC algorithm (Figure 10b) shows the least good comparison of the six, but, even then, the metrics are good (r^2 = 0.890; RMSD = 0.17). The envelope of points has a slight curve to it; correcting this using an empirical quadratic adjustment (see Table 2) leads to a slight improvement (Figure 10c).

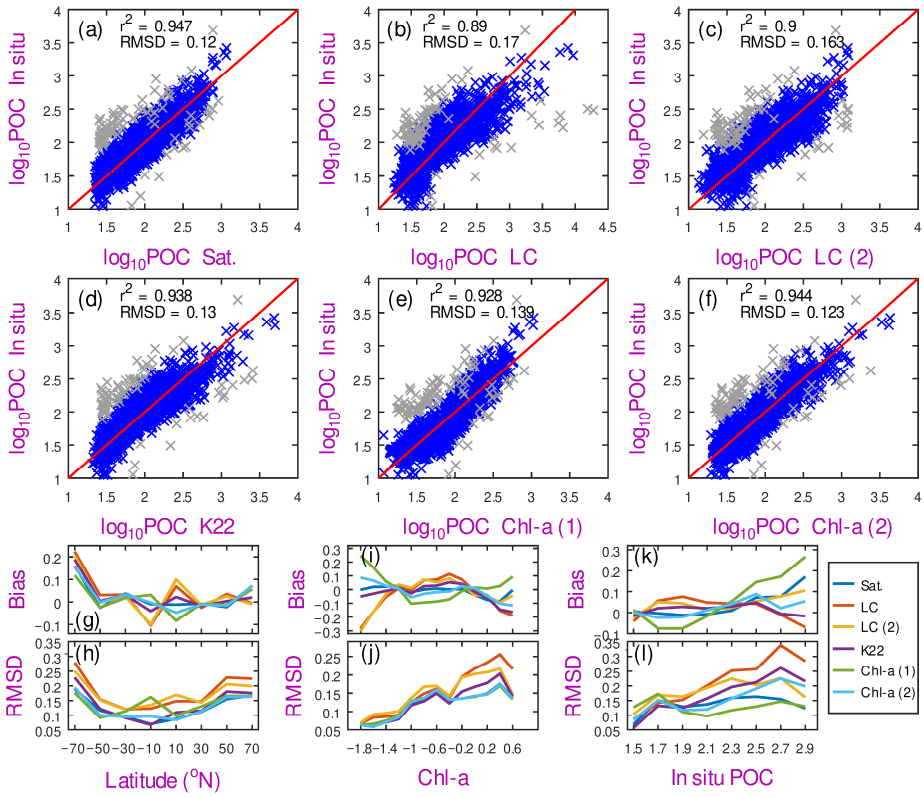

Figure 10. (**a**–**f**) Comparison of six satellite-based algorithms with matched in situ data. All algorithms have been adjusted so that the best-fitted straight line corresponds to $y = x$ (shown by the red lines). There are 3281 matchup points with valid satellite data, of which 82 (shown by grey symbols) were flagged as outliers in at least three comparisons, and are thus ignored in the fitting statistics. The bias (satellite-derived estimate minus in situ observation) and RMSD were then assessed as a function of (**g**,**h**) latitude, (**i**,**j**) the in situ value for Chl-a, and (**k**,**l**) the in situ POC value. [The "Sat" algorithm in panel (**a**) refers to the modified Stramski of Equation (4), whereas the rest are evaluations using BGC-Argo measurements. In this plot all POC and Chl-a values are displayed as $\log_{10}$ of the estimate in mg m^{-3}].

Table 2. Rescaled versions of previously published algorithms that best match our in situ dataset. [All algorithms are expressed in terms of logarithms].

Stramski	P = −1.5664 R +2.4949
LC (linear adjustment)	P = 1.3083 B + 5.7582 (=1.3083 G + 0.2843)
LC (quadratic adjustment)	P = −0.4216 B^2 − 1.0976 B + 2.3738
K22	P = 1.5885 K − 1.0581
Chlorophyll (linear adjustment)	P = 0.6203 C +2.355
Chlorophyll (quadratic adjustment)	P = 0.1863 C^2 + 0.8540 C + 2.2396

P is $\log_{10}$(POC estimate); R is $\log_{10}(R_{rs}(490)/R_{rs}(560))$; G is $\log_{10}$(POC estimate derived using coefficients of Galí et al., [7]); K is $\log_{10}$(POC estimate from K22); C is $\log_{10}$(Chl-a/mg m^{-3}); and B is $\log_{10}(b_{bp}(700)/\text{mg m}^{-3})$, with both Chl-a and $b_{bp}(700)$ being values calculated from the OC-CCI dataset, with the extrapolation for $b_{bp}(700)$ being $B = \log_{10}(b_{bp}(665)) + \chi\,[\log_{10}(b_{bp}(665)) - \log_{10}(b_{bp}(560))]$, where $\chi = [\log_{10}(700) - \log_{10}(560)]/[\log_{10}(665) - \log_{10}(560)]$.

However, the K22 algorithm (Figure 10d), incorporating the use of both b_{bp} and Chl-a estimates, produces better results, close to those of the modified Stramski. POC, like many measures of biological presence or activity, tends to have a strong connection with carbon

or chlorophyll biomass in phytoplankton. A logical extension is to assess how well POC can be determined from just satellite estimates of Chl-a. A simple linear relation between $\log_{10}$POC and $\log_{10}$Chl-a shows a slightly curved envelope (Figure 10e), which can be improved further using a quadratic adjustment (Figure 10f, see Table 2). This latter result is almost on par with the Stramski algorithm (Figure 10a).

The lower panels show how the anomalies (satellite-derived estimate minus in situ value) are partitioned as functions of latitude, Chl-a and the in situ POC value. One of the most striking results is that all the satellite algorithms (whether based on b_{bp}, Chl-a or ratio of reflectances) overestimate the in situ by 0.12 to 0.22 in the southernmost band (corresponding to overestimation by 32–66% in absolute units). As all these estimates are based on satellite-derived measures, this finding is independent of Argo floats with high Chl-a estimates in this region. There is much greater scatter at high Chl-a ($\log_{10}$Chl > 0) and high POC, which may indicate greater spatial and temporal variability in the in situ POC field. Across these three partitions, the Stramski algorithm and the quadratically adjusted chlorophyll-based algorithm show the lowest range of biases and scatter. Therefore, to produce consistent surface POC estimates from BGC-Argo observations we recommend a regional adjustment of the Chl-a values to match the CCI record and then a quadratic adjustment will allow a strong relationship between $\log_{10}$Chl and $\log_{10}$POC.

This paper has used linear rescaling (in logarithm space) multiple times in order to force algorithms to be unbiased with respect to each other and thus to allow a clearer investigation of the spatial and temporal differences between POC estimates. Considering the in situ dataset to be an unbiased record of the true values, the adjusted algorithms shown in Figure 10 are detailed in Table 2. Our linear result for b_{bp} alone returns slightly higher values than the implementation of LC by Gali et al. [7], and is also greater than the coefficients noted by Johnson et al. [18] in a focused study of Southern Ocean data only. The linear formulation in $\log_{10}$Chl-a is close to that noted by Legendre and Michard [19] and Stramski et al. [20]; our much larger in situ database allowed us to generate a robust quadratic adjustment that brought performance up to the standard of the best satellite-based algorithms.

4. Discussion

In this work we have determined POC algorithms for both satellite ocean colour (CCI) and BGC-Argo data that produce consistent results between the two platforms and correspond well with the limited in situ matchups. In identifying a consistent set of algorithms, a key issue is the definition of error that is used. Here, we have used fractional or relative errors on the grounds that a given percentage error at low or high POC concentrations is equally important. Therefore, practically all analyses were carried out in log space. This is particularly useful for looking at patterns of POC and associated temporal changes; if the chief concern is the POC budget for large basins then it may be more appropriate to work with absolute errors and achieve the best match possible for the largest concentrations.

Stramski et al. [20] have further developed satellite algorithms for estimating POC using the polynomials of logarithmic terms, including a hybrid version designed to work better in waters with very low POC (<25 mg m^{-3}). However, we did not evaluate those, as they involved separate tuning for each individual ocean colour sensor, rather than using the homogenous output from the ocean colour CCII programme.

A clear pattern throughout all the analyses was that the biological communities and their optical properties seemed to be different south of 45°S [21–23]. The initial discrepancies noted for POC appear to be related to the overestimation of Chl-a by the fluorescence sensors on BGC-Argo floats. Roesler et al. [2] had shown regional variations in the sensors' overestimation, with the factor being largest for parts of the Southern Ocean.

Our comparison with OC-CCI chlorophyll estimates confirmed this observation, and we established and used the bias to adjust the BGC-Argo values to match more consistently the satellite-derived values. Our comparison was simple, in that we just defined three separate regions for assessment, but as the region south of 45°S only had 4718 matchups with valid OC-CCI data, any separation into smaller regions could have been problematic. Indeed, it may be more appropriate to compartmentalise the algorithm for different biogeochemical provinces, for different seasons, or for different optical water types. However, that would still leave some classes with very few data for validation, especially in shallow coastal zones, given that standard BGC-Argo floats use a parking depth of 1000 m.

We assessed two algorithms for inferring POC from BGC-Argo float measurements. We note that Galí et al. [7] implemented the LC algorithm using different coefficients for the surface mixed layer in different regions; we had to use one set universally in order to have an algorithm that could be applied globally. Their implementation would have returned larger values in the subtropical gyres and lower ones in the Southern Ocean than those given by our "universal" algorithm. The LC algorithm only uses $b_{bp}(700)$ as input, and can occasionally results in unrealistically large values in regions with low Chl-a (see Figure 5). The approach, advocated by Koestner et al. [8], incorporates Chl-a, which, in effect, limits the estimate of POC in low chlorophyll waters. This modifies many of the outlier values, but slightly increases sensitivity to the different nature of waters south of 45°S.

There is significant disparity between the estimates from the chosen satellite POC algorithm based on the work by Stramski et al. [13] and applying the BGC-Argo-based algorithms to the satellite data (Figure 8g,h). From a practical perspective, it would make sense when aiming to create a consistent multi-sensor dataset to apply the same algorithm to both types of data. The Stramski algorithm cannot be applied to data from BGC-Argo floats, as these only provide measurements of Chl-a and $b_{bp}(700)$. Figure 8e,f show that the best agreement between satellite and Argo estimates of POC occur when identical algorithms are used on both sources.

The analysis in Figure 10 shows that a POC algorithm based solely on the estimates of Chl-a exhibits a performance almost equivalent to that of a dedicated satellite multi-wavelength algorithm developed by Stramski et al. [13]. For the data we used, we could identify no conditions for which the performance of the Chl-a POC algorithm was markedly worse than the Stramski algorithm. Admittedly, the comparison excluded coastal waters, which may encompass conditions where there are other inputs to the POC pool. However, BGC-Argo floats are not generally designed to operate in such shallow waters. Therefore, a simple algorithm based solely on Chl-a may be sufficient for producing POC estimates that are consistent between satellite and BGC-Argo observation.

5. Summary and Conclusions

In this paper we have considered various observables associated with Particulate Organic Carbon (POC) in the ocean from three platforms (satellites, BGC-Argo floats and ships), in order to assess how this key component of the ocean's biological carbon pool might best be monitored. The OC-CCI project has homogenised the data from over 25 years of continuous satellite observations from multiple sensors to produce self-consistent datasets of sea surface reflectances, optical backscatter, and associated derived variables. More recently the fleet of Argo floats has been enhanced with both backscatter and fluorescence sensors, although the coverage is, at present, far from uniform. These two datasets can augment the limited number of direct observations of POC by research vessels. Our work has focused on achieving the best agreement of these three sources for surface waters. There are a number of well-established algorithms for the estimation

of POC from satellite data but only two, to our knowledge [7,8] have been used with BGC-Argo measurements in a near-global context.

As there are few BGC-Argo profiles synoptic with ship-based in situ measurements, our approach has been to achieve a consistency between BGC-Argo and satellite estimates (see Figure 11), and then to use the greater number of ship–satellite matchups to assess different remote sensing approaches. The initial attempts to compare satellite and BGC-Argo POC estimates showed a clear regional bias, with the values from the floats being higher than satellite ones for the region south of 45°S. This could partly be attributed to the overestimation of Chl-a by BGC-Argo floats in these regions [2]. A regional adjustment of the floats' Chl-a and b_{bp} measurements then led to better consistency between satellite and Argo POC estimates (Figures 8c,d and 9). Very recently, the OneArgo project has developed a look-up table [24] for the correction of Chl-a values for their different physiological situations, so future releases of Chl-a data from floats will have a slightly different regional variation, and the empirical tuning that we used may no longer be needed.

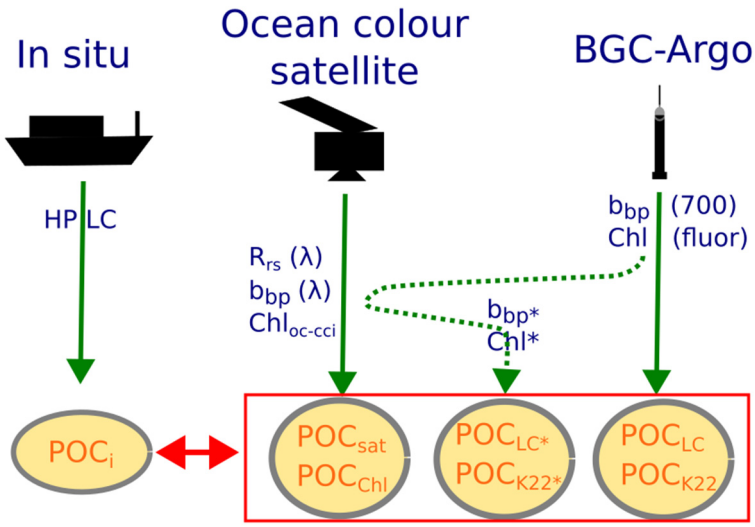

Figure 11. Summary of the various POC products compared. The in situ measure, POCi, comes from HPLC or other in vivo measurements. The OC-CCI project has homogenised the data from a number of ocean colour satellites to produce consistent records of remote sensing reflectance (R_{rs}), particulate backscatter (b_{bp}) and Chl-a (from the ratio of R_{rs} values). We evaluated three published satellite POC algorithms, and found that a modification of Stramski's second algorithm (to now use 560 nm) produced the best results; this is designated as POC_{sat}. An empirical algorithm, POC_{Chl}, based on the satellite records of Chl-a, was also developed. Two POC algorithms based on BGC-Argo data were also evaluated: POC_{LC} represents an estimate based solely on the floats' b_{bp}(700) data, and applying a universal set of coefficients from Galí et al. [7], whilst POC_{K22} represents the algorithm of Koestner et al. [8] that utilises both b_{bp}(700) and Chl-a from the BGC-Argo floats. A further pair of products, POC_{LC*} and POC_{K22*}, were developed to account for the overestimation of POC in the Southern Ocean; these involved an empirical adjustment of the b_{bp}(700) and Chl-a records of the Argo floats (indicated by asterisks) and then the application of the LC and K22 algorithms respectively. The BGC-Argo based algorithms were assessed against POC_{sat} for the ~36,000 matchups of floats with satellite data, and then rescaled versions of those algorithms were applied to satellite-derived estimates of b_{bp}(700) and Chl-a for comparison to the 3199 satellite matchups with in situ data.

Different algorithms applied to the same sets of satellite data led to RMSD values of ~0.10 (Figure 8g,h, corresponding to 27% in absolute units) whereas the same algorithm applied to the different sensors (satellite versus BGC-Argo) yielded an RMSD of only ~0.06 (Figure 8e,f, 15% in absolute units). However, all the algorithms evaluated showed a bias of ~0.12 to 0.22 in the southernmost region (Figure 10g), corresponding to overestimates of 32–66%.

There is a strong association between the recorded in situ POC values and all the bio-optical estimates, whether they are based on a ratio of sea surface reflectances, a b_{bp} estimate, a Chl-a estimate or a combination of the latter two. The modified Stramski algorithm (based on the ratio of reflectances at 560 nm and 490 nm) has the best correlation with the in situ POC values (RMSD = 0.120 in logarithmic units, corresponding to agreement to within 32% in absolute units), but a very similar level of performance is found for an algorithm based solely on the Chl-a estimate, which allows easy portability between satellite and BGC-Argo data. That algorithm results in an RMSD of 0.123 (equivalent to 33% in absolute units), with minimal variation with latitude or Chl-a value.

A POC dataset based solely on satellite data will tend to overestimate the amount of POC in very oligotrophic waters, because the standard algorithms cannot produce reliable estimates below $\log_{10}$POC = 1.35 (see Figure 4b). This would lead to an overestimate of the total POC in such areas. It is anticipated that a dataset that carefully melds data from satellites, ships and BGC-Argo floats could overcome this bias.

In the analyses covered in this paper, the various comparisons contained a small number of "outliers", which could be due to errors in the in situ data or in the satellite/BGC-Argo data (or simply a result of strong spatial inhomogeneity). To fully implement these POC algorithms for satellite or BGC-Argo data, further work is needed to assess the data quality from these sources, so their estimates can be trusted in the absence of in situ verification. Furthermore, to utilise the full potential of the BGC-Argo data, these algorithms need to be validated at depth. Both Galí et al. [7] and Koestner et al. [8] proposed algorithms to recover vertical profiles of POC, but the in situ data to provide independent validation appear very limited in their extent.

Author Contributions: Conceptualization, G.D.Q. and S.S.; methodology, G.D.Q.; software, G.D.Q.; formal analysis G.D.Q.; writing—original draft preparation, G.D.Q.; writing—review and editing, G.D.Q., S.S. and M.G.; and funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the UK National Centre for Earth Observation (NCEO), the Simons Foundation Project Collaboration on Computational Biogeochemical Modelling of Marine Ecosystems (CBIOMES, 549947, SS). MG acknowledges grant CEX2024-001494-S funded by AEI 10.13039/501100011033.

Data Availability Statement: There were three sources of data used: the BGC-Argo data were downloaded from Coriolis, the Ocean Colour CCI data are available at https://www.oceancolour.org/ (accessed 25 April 2024), and the in situ database used was as produced by Kong et al. [10]. We are grateful to the International Argo Programme [25,26] for collecting the data and making them freely available. Analysis was performed using Octave v6.4.0.; Matlab code for calculating Model 2 regressions is available [27], with that page also providing pointers to the background literature.

Acknowledgments: The authors are grateful to Gemma Kulk for a reprocessing of the matchup database, and to Giorgio Dall'Olmo and Nathan Briggs for discussions on regional biases in algorithms.

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

Appendix A. Application of Smoothing to Increase Satellite Matchups

To achieve a good statistical evaluation of a comparison between two different sensors, it is crucial to have a large number of matchups and that they cover the wide range of conditions that are encountered. This includes not just high and low values for the variable of interest, but (in our case), a span of latitudes, biogeochemical provinces, seasons, temperatures and cloud conditions. Also, as physical features, such as eddies, can affect thermal stability, stratification and phytoplankton community structure [28,29], matchups

should encompass cyclonic eddies, anticyclonic eddies and the regions between eddies. Our database contains 89,110 BGC-Argo profiles; of these, 22,487 had an exact match (same day and 4 km × 4 km pixel) with Chl-a or $b_{bp}(700)$ from the Ocean Colour CCI v6.0. To increase this number, we considered all CCI values within ±2 days and ±1 pixel, and calculated the median of these 45 values (to mitigate the effect of outliers); this resulted in 51,500 (51,495) matchups for Chl-a (b_{bp}).

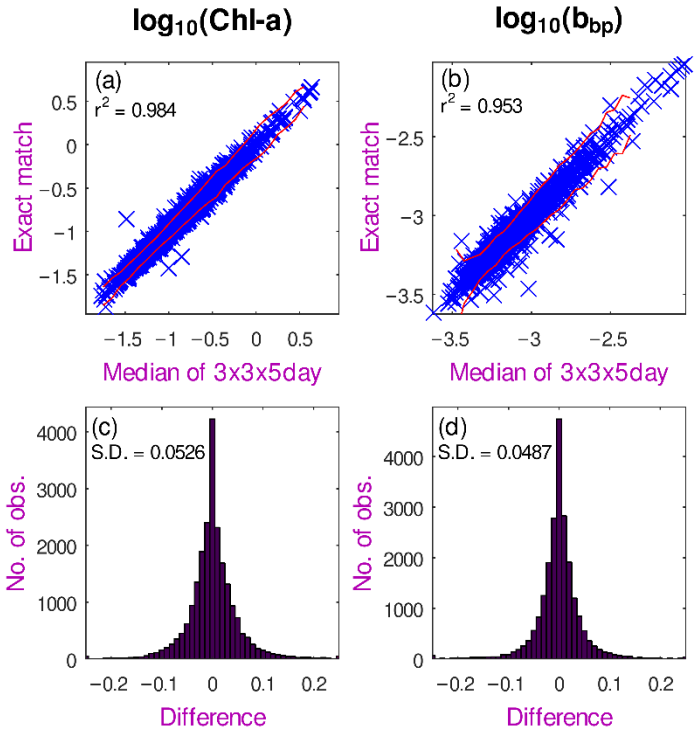

Figure A1. Comparison of CCI exact match data and median filtered values at the locations of BGC-Argo measurements. Scatter plots of (**a**) chlorophyll concentration and (**b**) $b_{bp}(700)$, with only a subset of points being drawn. The red lines indicate ±2 std. dev. about the mean. (**c**,**d**) Histograms of the differences between the two measures. (Note that because this comparison was performed for matchup locations with BGC-Argo floats, it does not encompass many shallow sea observations and is thus not representative of coastal upwelling regimes).

A comparison of the exact matchups with the calculated median representatives is shown in Figure A1. In many cases the median is identical to the exact match; in a few cases there is a wide disparity, which most likely indicates that the exact match had an anomalous record. The high correlation coefficients show that the use of this spatiotemporal median, which has more than doubled the number of matchups, does not appreciably add to the error budget. For both variables, the scatter in the plot (std. dev. shown by red lines) is generally ~75% greater for the largest values than for the lowest, indicating greater heterogeneity in conditions of high values of Chl-a and b_{bp}. (There is also increased variability in the lowest bin for Chl-a, which is likely due to the sensitivity of the inversion algorithm to measurement noise or atmospheric correction.)

Appendix B. Comparison of Satellite and In Situ Chlorophyll and Backscatter Estimates

In comparing POC estimates from satellite and BGC-Argo, it became apparent that a part of the difference originated in their different estimates of intermediate products, i.e., $b_{bp}(700)$ and Chl-a. The backscatter value at 700 nm is not a standard Ocean Colour CCI product; however, the estimates for the longer wavelengths correspond to a linear drop off

in logarithmic space, and so the values at 560 and 665 nm may be extrapolated to 700 nm in a consistent manner (see Figure 3). This follows the methodology used in the creation of the Ocean Colour CCI product that applies an exponential tail to define backscatter values at the selected frequencies (see Equation (6) of [30]).

We contrast the b_{bp} and Chl-a estimates from BGC-Argo with the values from the CCI dataset, with a focus on what adjustments are required to make the two datasets consistent. First, the scatter plots of both the b_{bp} and Chl-a data are used together to iteratively remove all the outliers. This reduced the initial matchup dataset from 37,870 members to 36,320. Some of these, with high BGC-Argo values and low satellite records, are presumed to be due to ingress of water into the floats' sensors when at depth, and often have full profiles that are awry. For many floats, all successive profiles after a certain date will be unusable. In our analysis, we had not implemented the quality flags in the Argo products; had we done so, it is likely that many of these would have been discarded. Further outliers may be due to unrecognised clouds or sea ice affecting the satellite-measured sea surface reflectances and causing errors in the CCI values. Our simple data editing procedure to iteratively remove points more than three S.D. from the mean relationship only removes 4% of data pairs and deals with both sensor problems and poor matchups of satellite and BGC-Argo observations.

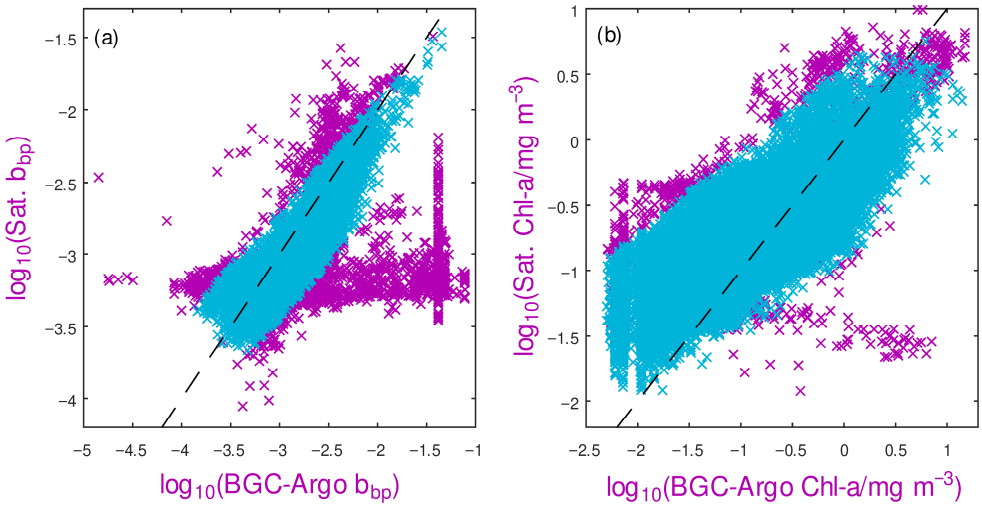

Figure A2. Matchups of all BGC-Argo and CCI estimates of (**a**) b_{bp} and (**b**) Chl-a, with black dashed lines indicating $y = x$. Initial data selection discarded in situ measurements of $\log_{10}(b_{bp})$ less than -5 and $\log_{10}$(Chl-a) less than -2.3. [Note that satellite retrievals of $\log_{10}$(Chl-a) are not credible below -2, and indeed in situ analytical techniques are not reliable below 0.01 mg m^{-3}, so there is no supporting validation for values that low.] An iterative technique is used to remove the extreme outliers (shown in purple) from each plot in turn leaving the good matchups in blue.

As Roesler et al. [2] had identified regional biases in the BGC-Argo estimates of Chl-a (with the Southern Ocean being particularly anomalous), we segregated our data into three broad regions (90–45°S, 45°S–45°N and 45–90°N). A direct comparison of the b_{bp} values (Figure A3a,b) shows an almost linear relationship (in log space), with little difference between the three regions. The satellite algorithm does not produce values below -3.7, whereas the Argo sensors do produce a small number of lower values; consequently, the relationship between the two appears to level out at low values. Figure A3c shows that the bias between the satellite and BGC-Argo estimates of b_{bp} has little dependency on Chl-a.

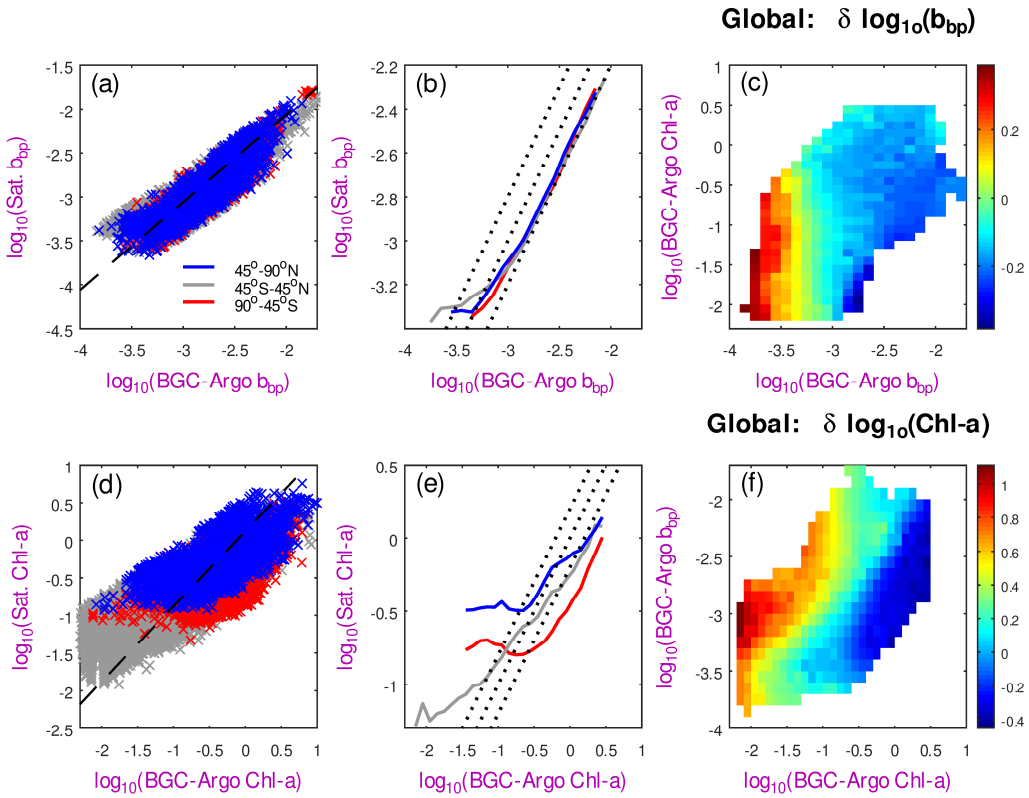

Figure A3. Comparison of BGC-Argo and satellite (Ocean Colour CCI) estimates of $b_{bp}(700)$ and Chl-a. (**a**) Scatter plot of b_{bp} values, coloured by region, with dashed line showing a simple addition of -0.04. (**b**) Curves showing the mean satellite values for the three regions, binned in steps of 0.1, and the dotted lines show a slope of one, with offsets of ± 0.2. (There are ten times as many observations in the middle region as either of the high-latitude ones; hence, its curve can be reliably calculated over a greater range of x-values.) (**c**) The mean bias of satellite minus Argo as a function of both b_{bp} and Chl-a (using all data together). (**d**–**f**) Corresponding plots for Chl-a, with dashed line in (**d**) showing an offset of 0.17 and the dotted lines in (**e**) having offsets of ± 0.2. Note that the discrepancy in the Chl-a estimates does vary appreciably with b_{bp}.

The picture is rather different for Chl-a, with a marked deviation from a straight line and clear differences between the regions (Figure A3d,e). The fluorometers on the BGC-Argo floats can yield very weak values for fluorescence, which are used to infer very low chlorophyll concentrations. (Indeed, after the necessary step of removing the noise floor of the fluorescence values, some returned estimates are negative, which is clearly meaningless.) In situ analysis using HPLC rarely delivers estimates below 0.01 mg m^{-3}, and the unrealistically low BGC-Argo values are associated with a wide range of satellite CCI estimates, averaging to $\log_{10}$(Chl-a) values a little below -1.

At $\log_{10}$(Chl-a) values greater than zero, the slope of the relationship between BGC-Argo's fluorometric estimates and CCI's radiometric estimates is close to one. The satellite values are then typically lower by 0.2 in logarithmic units, but for the region south of 45°S (shown by the red lines), the satellite values are a further 0.2 units lower, or equivalently the BGC-Argo values for this southern sector are 0.2 units higher than for the rest of the globe, corresponding to a multiplicative factor of 1.6 in absolute units. Even greater disparities are noted for records in the southern region within the realm $-0.7 < \log_{10}\text{Chl-a} < 0$. Figure 2 of Roesler et al. [2] showed fluorescence-derived values in the Atlantic, Pacific and Indian sectors of the Southern Ocean to be ~2.5 times greater than for the other sectors, but with a wide range of values.

To create a homogeneous series of b_{bp} and Chl-a values spanning BGC-Argo and satellite observations, the measurements from one sensor need to be adjusted to match

those from the other. Here, we have chosen to adjust the Argo observations to match the satellite ones. At the most basic level, this can be achieved by a simple offset to be added to all values. For the log values of b_{bp} and Chl-a these biases are −0.04 and 0.17 respectively when evaluated on a global basis; after this, adjustment the RMSD values are 0.137 and 0.373 (see Tables A1 and A2). Figure A3e shows that for Chl-a, there are different offsets for our three specified regions. (Note that 74% of the matchups occur in the region spanning 45°S–45°N, so the accuracy of its adjustment will always dominate the three-region comparison.)

Table A1. Root mean square discrepancies between BGC-Argo observations and satellite-inferred values of $\log_{10}(b_{bp})$ for locations of matchups, upon application of various empirical adjustments.

b_{bp} Error	Simple Offset	Line-Fitted	f(b_{bp}_Argo)	f(b_{bp},Chl_Argo)
Global	0.137	0.103	0.091	0.088
90–45°S	0.100	0.094	0.092	0.088
45°S–45°N	0.135	0.096	0.084	0.081
45–90°N	0.136	0.122	0.116	0.109
3 regions	0.132	0.099	0.088	0.085
Area weighted	0.130	0.098	0.088	0.084

There are 4718 matchups in the southern region, 26,988 in the central region and 4524 in the northern one. "Simple offset" corresponds to the bias indicated by the dashed line in Figure A3a; the next column is for a linear adjustment, the subsequent one is for binned representation in Figure A3b, and the last column represents correction by the function of both b_{bp} and Chl-a shown in Figure A3c. The results in the 'global' row are for a single adjustment applied to all the data, with the subsequent three rows giving the error in three broad latitudinal bands when separate adjustments are defined. The penultimate row shows the overall global error if the three regions are separately adjusted, with the last row showing the effective error if the results for the three regions are weighted according to the area within them that is at least 1000 m deep (the typical parking depth of Argo floats).

Table A2. Same as Table A1 but for $\log_{10}$(Chl-a).

Chl-a Error	Simple Offset	Line-Fitted	f(Chl_Argo)	f(b_{bp},Chl_Argo)
Global	0.373	0.233	0.230	0.187
90–45°S	0.290	0.199	0.167	0.151
45°S–45°N	0.339	0.215	0.211	0.166
45–90°N	0.363	0.222	0.211	0.159
3 regions	0.337	0.215	0.207	0.164
Area weighted	0.332	0.213	0.204	0.163

Improved consistency can be achieved with more complex adjustments. Using a linear fit (as opposed to a simple offset) reduces the RMSD by about 25% for both b_{bp} and Chl-a. Incorporating more detailed adjustments of BGC-Argo values that are functions of both b_{bp} and Chl-a (see Figure A3c,f) only makes a small further improvement to b_{bp} (as that correction was nearly linear), but has a greater impact on the Chl-a values.

References

1. Volk, T.; Hoffert, M.I. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. In *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present*; Sundquist, E.T., Broecker, W.S., Eds.; Wiley: Hoboken, NJ, USA, 1985; pp. 99–110. [CrossRef]
2. Roesler, C.; Uitz, J.; Claustre, H.; Boss, E.; Xing, X.; Organelli, E.; Briggs, N.; Bricaud, A.; Schmechtig, C.; Poteau, A.; et al. Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. *Limnol. Oceanogr. Meth.* **2017**, *15*, 572–585. [CrossRef]
3. Ciotti, Á.; Lewis, M.R.; Cullen, J.J. Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnol. Oceanogr.* **2002**, *2*, 404–417. [CrossRef]
4. Uitz, J.; Stramski, D.; Reynolds, R.A.; Dubranna, J. Assessing phytoplankton community composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-sensing reflectance in open-ocean environments. *Remote. Sens. Environ.* **2015**, *171*, 58–74. [CrossRef]
5. Pérez, G.L.; Galí, M.; Royer, S.-J.; Gerea, M.; Ortega-Retuerta, E.; Gasol, J.M.; Marrasé, C.; Simó, R. Variability of phytoplankton light absorption in stratified waters of the NW Mediterranean Sea: The interplay between pigment composition and the packaging effect. *Deep Sea Res. I* **2021**, *169*, 103460. [CrossRef]
6. Evers-King, H.; Martinez-Vicente, V.; Brewin, R.J.W.; Dall'Olmo, G.; Hickman, A.E.; Jackson, T.; Kostadinov, T.S.; Krasemann, H.; Loisel, H.; Röttgers, R.; et al. Validation and intercomparison of ocean color algorithms for estimating particulate organic carbon in the oceans. *Front. Mar. Sci.* **2017**, *4*, 251. [CrossRef]
7. Galí, M.; Falls, M.; Claustre, H.; Aumont, O.; Bernardello, R. Bridging the gaps between particulate backscattering measurements and modeled particulate organic carbon in the ocean. *Biogeosciences* **2022**, *19*, 1245–1275. [CrossRef]
8. Koestner, D.; Stramski, D.; Reynolds, R.A. A multivariable empirical algorithm for estimating particulate organic carbon concentration in marine environments from optical backscattering and chlorophyll-a measurements. *Front. Mar. Sci.* **2022**, *9*, 941850. [CrossRef]
9. Thierry, V.; Claustre, H.; de Fommervault, O.P.; Zilberman, N.; Johnson, K.S.; King, B.A.; Wijffels, S.E.; Bhaskar, U.T.V.S.; Balmaseda, M.A.; Belbeoch, M.; et al. Advancing ocean monitoring and knowledge for societal benefit: The urgency to expand Argo to OneArgo by 2030. *Front. Mar. Sci.* **2025**, *12*, 1593904. [CrossRef]
10. Kong, C.E.; Sathyendranath, S.; Jackson, T.; Stramski, D.; Brewin, R.J.W.; Kulk, G.; Jönsson, B.F.; Loisel, H.; Galí, M.; Le, C. Comparison of ocean-colour algorithms for particulate organic carbon in global ocean. *Front. Mar. Sci.* **2024**, *11*, 1309050. [CrossRef]
11. Sathyendranath, S.; Brewin, R.J.W.; Brockmann, C.; Brotas, V.; Calton, B.; Chuprin, A.; Cipollini, P.; Couto, A.B.; Dingle, J.; Doerffer, R.; et al. An ocean-colour time series for use in climate studies: The experience of the Ocean-Colour Climate Change Initiative (OC-CCI). *Sensors* **2019**, *19*, 4285. [CrossRef]
12. Loisel, H.; Nicolas, J.-M.; Deschamps, P.-Y.; Frouin, R. Seasonal and inter-annual variability of particulate organic matter in the global ocean. *Geophys. Res. Lett.* **2002**, *29*, 2196. [CrossRef]
13. Stramski, D.; Reynolds, R.A.; Babin, M.; Kaczmarek, S.; Lewis, M.R.; Röttgers, R.; Sciandra, A.; Stramska, M.; Twardowski, M.S.; Franz, B.A.; et al. Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. *Biogeosciences* **2008**, *5*, 171–201. [CrossRef]
14. Loisel, H.; Bosc, E.; D. Stramski, D.; Oubelkheir, K.; Deschamps, P.-Y. Seasonal variability of the backscattering coefficient in the Mediterranean Sea based on satellite SeaWiFS imagery. *Geophys. Res. Lett.* **2001**, *28*, 4203–4206. [CrossRef]
15. Cetinić, I.; Perry, M.J.; Briggs, N.T.; Kallin, E.; D'Asaro, E.A.; Lee, C.M. Particulate organic carbon and inherent optical properties during 2008 North Atlantic Bloom Experiment. *J. Geophys. Res.* **2012**, *117*, C06028. [CrossRef]
16. Galí, M.; Orihuela-García, A.; Ruprich-Robert, Y.; Lapin, V.; Sánchez-Urrea, M.; Fontela, M.; Llort, J.; Sicardi, V.; Bernardello, R. Convection injects labile particulate organic carbon to the deep ocean. *ESS Open Arch.* **2025**. [CrossRef]
17. Loveday, B.R.; Smyth, T.J. A 40-year global data set of visible-channel remote-sensing reflectances and coccolithophore bloom occurrence derived from the Advanced Very High Resolution Radiometer catalogue. *Earth Sys. Sci. Data* **2018**, *4*, 2043–2054. [CrossRef]
18. Johnson, K.S.; Plant, J.N.; Coletti, L.J.; Jannasch, H.W.; Sakamoto, C.M.; Riser, S.C.; Swift, D.D.; Williams, N.L.; Boss, E.; Haëntjens, N.; et al. Biogeochemical sensor performance in the SOCCOM profiling float array. *J. Geophys. Res. Oceans* **2017**, *122*, 6416–6436. [CrossRef]
19. Legendre, L.; Michaud, J. Chlorophyll a to estimate the particulate organic carbon available as food to large zooplankton in the euphotic zone of oceans. *J. Plankton Res.* **1999**, *21*, 2067–2083. [CrossRef]
20. Stramski, D.; Joshi, I.; Reynolds, R.A. Ocean color algorithms to estimate the concentration of particulate organic carbon in surface waters of the global ocean in support of a long-term data record from multiple satellite missions . *Remote Sens. Environ.* **2022**, *269*, 112776. [CrossRef]

21. Johnson, R.; Strutton, P.G.; Wright, S.W.; McMinn, A.; Meiners, K.M. Three improved satellite chlorophyll algorithms for the Southern Ocean. *J. Geophys. Res. Oceans* **2013**, *118*, 3694–3703. [CrossRef]
22. Haëntjens, N.; Boss, E.; Talley, L.D. Revisiting ocean color algorithms for chlorophyll a and particulate organic carbon in the Southern Ocean using biogeochemical floats. *J. Geophys. Res. Oceans* **2017**, *122*, 6583–6593. [CrossRef]
23. Moutier, W.; Thomalla, S.J.; Bernard, S.; Wind, G.; Ryan-Keogh, T.J.; Smith, M.E. Evaluation of chlorophyll-a and POC MODIS Aqua products in the Southern Ocean. *Remote Sens.* **2019**, *11*, 1793. [CrossRef]
24. Sauzède, R.; Schmechtig, C.; Renosh, P.R.; Uitz, J.; Claustre, H. Global look-up table of physiological ratios for the real-time adjustment of chlorophyll-a fluorescence within the OneArgo framework. *SEANOE* **2025**. [CrossRef]
25. Wong, A.P.S.; Wijffels, S.E.; Riser, S.C.; Pouliquen, S.; Hosoda, S.; Roemmich, D.; Gilson, J.; Johnson, G.C.; Martini, K.; Murphy, D.J.; et al. Argo data 1999–2019: Two million temperature-salinity profiles and subsurface velocity observations from a global array of profiling floats. *Front. Mar. Sci.* **2020**, *7*, 700. [CrossRef]
26. Schmechtig, C.; Wong, A.; Maurer, T.; Bittig, H.; Thierry, V. *Argo Quality Control Manual for Biogeochemical Data. v1.1*; Bio-Argo Group: Niterói, Brazil, 2023; Volume v1.1. [CrossRef]
27. Matlab. Matlab Scripts for Model I and Model II Regressions. A Brief History of Model II Regression Analysis. Available online: https://www.mbari.org/technology/matlab-scripts/linear-regressions/ (accessed on 3 March 2026).
28. Gaube, P.; McGillicuddy, D.J., Jr.; Chelton, D.B.; Behrenfeld, M.J.; Strutton, P.G. Regional variations in the influence of mesoscale eddies on near-surface chlorophyll. *J. Geophys. Res. Oceans* **2014**, *119*, 8195–8220. [CrossRef]
29. Han, G.; Quartly, G.D.; Chen, G.; Yang, J. Satellite-observed SST and chlorophyll reveal contrasting dynamical-biological effects of mesoscale eddies in the North Atlantic. *Environ. Res. Lett.* **2024**, *19*, 104035. [CrossRef]
30. IOCCG. Steps and Calculations of the Quasi-Analytical Algorithm (QAA_v6). Available online: https://ioccg.org/wp-content/uploads/2020/11/qaa_v6_202011.pdf (accessed on 24 February 2026).