Remote Sensing of Environment 286 (2023) 113417

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Performance assessment and validation of ocean color sensor-specific algorithms for estimating the concentration of particulate organic carbon in oceanic surface waters from satellite observations

Ishan D. Joshi [a,*], Dariusz Stramski [a], Rick A. Reynolds [a], Dale H. Robinson [b]

[a] Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0238, USA
[b] University of California Santa Cruz, affiliated with NOAA CoastWatch at NOAA Southwest Fisheries Science Center, Santa Cruz, CA 95060, USA

ARTICLE INFO

Edited by Dr. Menghua Wang

Keywords:
Satellite remote sensing
Ocean color
Particulate organic carbon
Global algorithms
Validation analysis

ABSTRACT

Global observations of ocean color from space provide a unique capability to conduct multi-decadal studies of biogeochemical constituents and processes in the upper ocean layers at the regional, basin, and global scales. The concentration of particulate organic carbon (POC) in the surface ocean waters is one of the biogeochemically important data products derivable from satellite ocean color observations. The data record of standard global POC product is available from NASA and has been obtained from the empirical power-function relationship between POC and a single blue-to-green band ratio of ocean remote-sensing reflectance, $R_{rs}(\lambda)$, which is here referred to as the BR-PF algorithm. Recently, we formulated a new suite of satellite ocean-color sensor-specific global POC algorithms that are referred to as hybrid algorithms because they contain two components, one based on the maximum band-ratio (MBR) and the other on the band ratio difference index (BRDI) of reflectance (Stramski et al., 2022). The present study describes the validation analysis and provides performance assessment of hybrid algorithms for SeaWiFS, MODIS, and VIIRS-SNPP satellite ocean color sensors. For comparison, the standard global (BR-PF) algorithms and the color-index (CI) algorithm from the literature were also analyzed. The analyses were made with an in situ validation dataset containing concurrent measurements of POC and $R_{rs}(\lambda)$ as well as three satellite validation datasets containing data matchups of field measurements of POC and satellite-derived $R_{rs}(\lambda)$ from the SeaWIFS, MODIS-Aqua, and VIIRS-SNPP missions. The hybrid algorithms showed an overall improvement of performance compared to other algorithms, in particular in waters with relatively high POC compared to the standard BR-PF algorithms. The aggregate bias of hybrid algorithm-derived POC is small or negligible (<10%) for the validation datasets, and a measure of median percentage difference is about 20% for the in situ validation dataset and ranges between about 20% and 30% for the satellite validation datasets. In addition, comparisons of POC retrievals from the three sensor-specific hybrid algorithms show good inter-sensor consistency. The validation results of this study indicate that hybrid algorithms have high potential to provide improved POC retrievals within a broader range of POC than the predecessor global algorithms and the capability for generating a consistent long-term POC data record from multiple satellite missions.

1. Introduction

Global observations of ocean color from space have been made continuously since the late 1990s and will continue into the foreseeable future owing to multiple satellite missions whose operational lifetimes overlap. This observational approach provides a unique capability to monitor long-term trends in satellite-derived data products characterizing biogeochemical constituents and processes in the upper ocean layers at the regional, basin, and global scales. One such data product of significant interest to studies of ocean carbon cycles and biogeochemistry is the concentration of particulate organic carbon (POC) in the surface ocean waters. POC is associated with the content of organic carbon in suspended particulate matter that includes various autotrophic and heterotrophic organisms (phytoplankton, bacteria, zooplankton), viruses, as well as organic detrital (non-living) particles. Compared to dissolved carbon pools in the ocean, the standing stock of POC is relatively small but is essential to the marine food web and responsible for relatively large carbon fluxes, including the carbon

* Corresponding author.
E-mail address: isjoshi@ucsd.edu (I.D. Joshi).

https://doi.org/10.1016/j.rse.2022.113417
Received 11 May 2022; Received in revised form 5 December 2022; Accepted 10 December 2022
Available online 3 January 2023

export from the surface into the deep ocean via sinking particles as part of the "biological pump", which provides a mechanism for sequestering carbon in the deep ocean and bottom sediments (e.g., Behrenfeld et al., 2005; Bianchi and Bauer, 2011; Boyd and Newton, 1999; Fenchel, 1988; Field et al., 1998; Longhurst and Harrison, 1989; Volk and Hoffert, 1985).

The use of satellite ocean color observations to estimate POC in surface waters was first demonstrated with spaceborne measurements made with SeaWiFS (Sea-viewing Wide Field-of-view Sensor) over the Southern Ocean (Stramski et al., 1999). That study used a combination of simultaneously collected in situ measurements of POC (in units of mg m^{-3}), the spectral particulate backscattering coefficient of seawater, $b_{bp}(\lambda)$ (in m^{-1}), and the spectral remote-sensing reflectance of the ocean, $R_{rs}(\lambda)$ (in sr^{-1} where λ is light wavelength in vacuum), to propose a two-step empirical algorithm in which the seawater inherent optical property of b_{bp} is derived first from R_{rs}, and then POC is estimated from b_{bp}. Since then, numerous approaches for ocean color POC algorithms have been explored for regional, basin-scale, or global ocean applications (Allison et al., 2010; Gardner et al., 2006; Le et al., 2018; Liu et al., 2021; Loisel et al., 2002; Mishonov et al., 2003; Pabi and Arrigo, 2006; Son et al., 2009; Stramska and Stramski, 2005; Stramski et al., 2008; Tran et al., 2019; Wang et al., 2020). The empirical algorithms based on a direct relationship between POC and $R_{rs}(\lambda)$ established from concurrent in situ measurements of these variables have emerged as the currently most satisfactory approach for applications to large-scale imagery of satellite-derived $R_{rs}(\lambda)$, especially global or basin-scale applications characterized by the dominance of open-ocean pelagic environments (Allison et al., 2010; Stramska and Stramski, 2005; Stramski et al., 2008). In particular, the relationships between POC and the blue-to-green reflectance band ratio were found to provide improved estimates of POC across a range of diverse water bodies compared to POC estimates obtained from algorithms that involve an intermediate step of retrieval of $b_{bp}(\lambda)$ from $R_{rs}(\lambda)$. This finding can be attributed largely to significant variations in the relationship between POC and $b_{bp}(\lambda)$ that are associated with potentially large changes in characteristics of particulate assemblages (e.g., composition and particle size distribution) across diverse aquatic environments (Allison et al., 2010; Cetinić et al., 2012; Rasse et al., 2017; Reynolds et al., 2016; Stramski et al., 1999; Stramski et al., 2008).

At present, the standard satellite-derived global POC data product generated by the NASA Ocean Biology Processing Group (OBPG) is based on an empirical algorithm that estimates POC from the blue-to-green reflectance band ratio using a power function, hereafter referred to as the BR-PF algorithm. This algorithm was formulated with data collected in open waters of the Pacific and Atlantic waters with surface POC ranging from about 10 to 300 mg m^{-3} (Stramski et al., 2008). This range represents most situations within the vast areas of open-ocean pelagic environments, although higher POC values can be encountered in open-ocean waters and more commonly in coastal regions. The standard global POC product from satellite missions equipped with multi-spectral ocean color sensors, in particular SeaWiFS on SeaStar satellite and MODIS (Moderate Resolution Imaging Spectroradiometer) on Aqua satellite, have been utilized to provide estimates of global and basin-scale distributions of surface POC and standing stock of POC integrated within the upper water column (Evers-King et al., 2017; Stramska and Cieszyńska, 2015). Using a large satellite-in situ matchup dataset, Evers-King et al. (2017) also presented validation analysis of POC algorithms, including the standard BR-PF algorithm, and demonstrated an overall satisfactory level of performance of the BR-PF algorithm with a very small aggregate bias in satellite-derived POC relative to measured POC, especially for the range within which the BR-PF algorithm was originally formulated. Importantly, however, there is evidence that the algorithms based on a single blue-to-green reflectance ratio such as BR-PF are likely to yield underestimated POC at relatively high POC (Evers-King et al., 2017; Le et al., 2018; Liu et al., 2021).

In a recent study we addressed a need to reexamine the formulation of satellite sensor-specific global POC algorithms for applications to different ocean color satellite missions in support of the generation of the long-term consistent data record of oceanic POC (Stramski et al., 2022, hereafter referred to as SJR2022). We note that the connotation "global" in connection with "algorithm" does not imply the intent of algorithm applicability to every type of water body within the global ocean. The connotation "global" is here used with an understanding that the algorithms are designed to optimize their performance over vast areas of open-ocean pelagic environments. In SJR2022, the hybrid algorithms that were formulated consisted of two components, the maximum band ratio (MBR) component and the band ratio difference index (BRDI) component. The BRDI component involves a difference of two band ratios of reflectance from the blue-green spectral region and is applied only in very clear waters when POC is very low (below approximately 20 mg m^{-3}). The MBR component utilizes three band ratios of reflectance from the blue-green spectral region and is applied over the entire remaining portion of the POC range. The SJR2022 study demonstrated the potential for improved POC estimates from the hybrid algorithm compared to the BR-PF algorithm, especially at higher POC above a few hundred of mg m^{-3} and at very low POC.

The sensor-specific hybrid algorithms can provide the next-generation suite of POC algorithms for the routine processing of satellite global observations from multiple missions, so it is important and timely to conduct a thorough validation analysis of these algorithms and also compare such validation results with the standard BR-PF algorithms. Such validation analyses and comparisons are the main objectives of the present study. To address these objectives, we assembled two distinct types of validation datasets: the in situ validation dataset and the satellite-in situ matchup validation dataset. The primary focus of our validation analyses is on the global hybrid algorithms for SeaWiFS, MODIS, and VIIRS-SNPP (Visible Infrared Imaging Radiometer Suite on Suomi National Polar Orbiting Partnership satellite) sensors and comparisons with corresponding sensor-specific standard BR-PF algorithms used currently by NASA OBPG. For additional comparative purposes, the validation results are also presented for another global POC algorithm that is based on the color index and described in Le et al. (2018).

2. Data and methods

2.1. In situ and satellite validation datasets for POC algorithm assessment and intercomparison analysis

Two validation datasets were assembled from public data sources including SeaWiFS Bio-optical Archive and Storage System, Biological and Chemical Oceanography Data Management Office, PANGAEA, National Center for Environmental Information, and individual project websites (e.g., U.S. Joint Global Ocean Flux Study and Atlantic Meridional Transect). These datasets are named based on the method used to determine the spectral remote-sensing reflectance, $R_{rs}(\lambda)$, i.e., in situ validation dataset and satellite validation dataset.

The method of field measurement of POC was considered as the primary inclusion criterion to generate the validation datasets of concurrent measurements of POC and $R_{rs}(\lambda)$. Specifically, our primary approach was to accept the POC determinations with a method similar to the JGOFS protocol, which includes the preparation of samples by filtration of discrete water samples through the GF/F filters and the analysis consisting of removal of particulate inorganic carbon by acid treatment of sample filters, high-temperature combustion and measurement with CHN analyzer, and the subtraction of carbon content on the blank filters (also called blank-correction) from the sample values for obtaining the final concentration of POC (Knap et al., 1996). This method of POC determination was used in the creation of the algorithm development dataset that underlies the formulation of the primary set of global hybrid POC algorithms presented in the SJR2022 study (see Table 4 in Stramski et al., 2022). We note that the blank-correction of the JGOFS POC protocol does not account for the DOC adsorption and the

potential loss of POC due to cell breakage from fragile plankton species or passage of some particles (mainly submicron) through the GF/F filter (nominal pore size of 0.7 μm) during sample water filtration, which are among several sources of uncertainty in POC measurements (Cetinić et al., 2012; Collos et al., 2014; Gardner et al., 2003; Turnewitsch et al., 2007). While efforts have been made to propose a correction for the positive bias caused by DOC adsorption (Moran et al., 1999; Novak et al., 2018), no standard methodology is available to date to correct for the negative bias associated with the losses of POC during filtration. Because the primary set of global POC algorithms developed in SJR2022 and other existing ocean color POC algorithms were formulated using the blank-corrected POC measurements without correction for these opposite biasing effects, the algorithm assessment and intercomparison analyses in this study were focused on such blank-corrected POC measurements made according to the JGOFS protocol.

Our main objective is to validate the global POC algorithms that are applicable mostly to vast pelagic areas of the world's oceans, therefore we did not include the measurements taken in optically shallow waters where light reflected off the seafloor contributes significantly to the water-leaving radiance. All stations of in situ and satellite validation datasets satisfied an optically deep-water criterion, i.e., the attenuation depth < water column depth (H_B). The attenuation depth was computed as $1.3/K_d(490)$, which represents the first attenuation depth, $1/K_d(490)$, increased by 30% (Gordon and McCluney, 1975). The diffuse attenuation coefficient for downwelling plane irradiance at 490 nm, $K_d(490)$, was estimated from the current NASA $K_d(490)$ algorithm (Morel et al., 2007), whereas bottom depth H_B was obtained from NOAA ETOPO 1 Arc-minute global relief model data (Amante and Eakins, 2009).

Only near-surface measurements of POC taken within the top 10 m layer of the ocean were considered in the generation of validation datasets. If POC measurements were made at multiple depths within the 10 m surface layer at a given oceanographic station, an average POC value from these measurements was used.

2.1.1. *In situ validation dataset*

The in situ validation dataset comprises two subsets of data, each corresponding to a different approach for measuring radiometric variables near the surface of the ocean. The in-water radiometry validation dataset (IWR) is based on underwater measurements of upwelling radiance accompanied with either above-water or underwater measurements of downwelling plane irradiance. The above-water radiometry validation dataset (AWR) dataset is based on above-water measurements of both the upwelling radiance and downwelling irradiance. To minimize the potential effects associated with differences in data processing and data quality control (QC) practices between the cruises, both in-water and above-water radiometric measurements for all cruises were reprocessed and quality-controlled with the same consistent methodology that is generally recommended in the literature (Austin, 1974; IOCCG Protocol Series, 2019; Mobley, 1999; Stramski et al., 2008). More details on in situ radiometric data collection, processing, and QC are provided in Supplementary Material.

In summary, the final IWR and AWR datasets consist of 174 and 335 coexisting pairs of near-surface POC and $R_{rs}(\lambda)$ measurements, respectively, which were collected from a broad range of ocean biogeochemical provinces with locations varying greatly in terms of water column depth across all major oceanic basins (Fig. 1). It is important to note that the in situ $R_{rs}(\lambda)$ determinations were not corrected for bidirectional effects but represent the actual measurements under given environmental conditions. The assembled in situ validation dataset is independent of data that were used to formulate the hybrid algorithms in SJR2022. The analysis of this independent dataset allows for the assessment of algorithm performance and characterization of uncertainties of POC estimation which are associated primarily with the algorithm formulation itself, although the uncertainties of in situ measurements of POC and $R_{rs}(\lambda)$ also have some effect on these validation results. The characterization of sources of uncertainties and typical levels of uncertainty in field determinations of POC and $R_{rs}(\lambda)$ are presented in IOCCG Protocol Series (2019) and IOCCG Protocol Series (2021). Details on research programs, cruises, number of stations, radiometric measurements, and basin-wide statistics for in situ validation datasets are given in Table S1 (Supplementary Material).

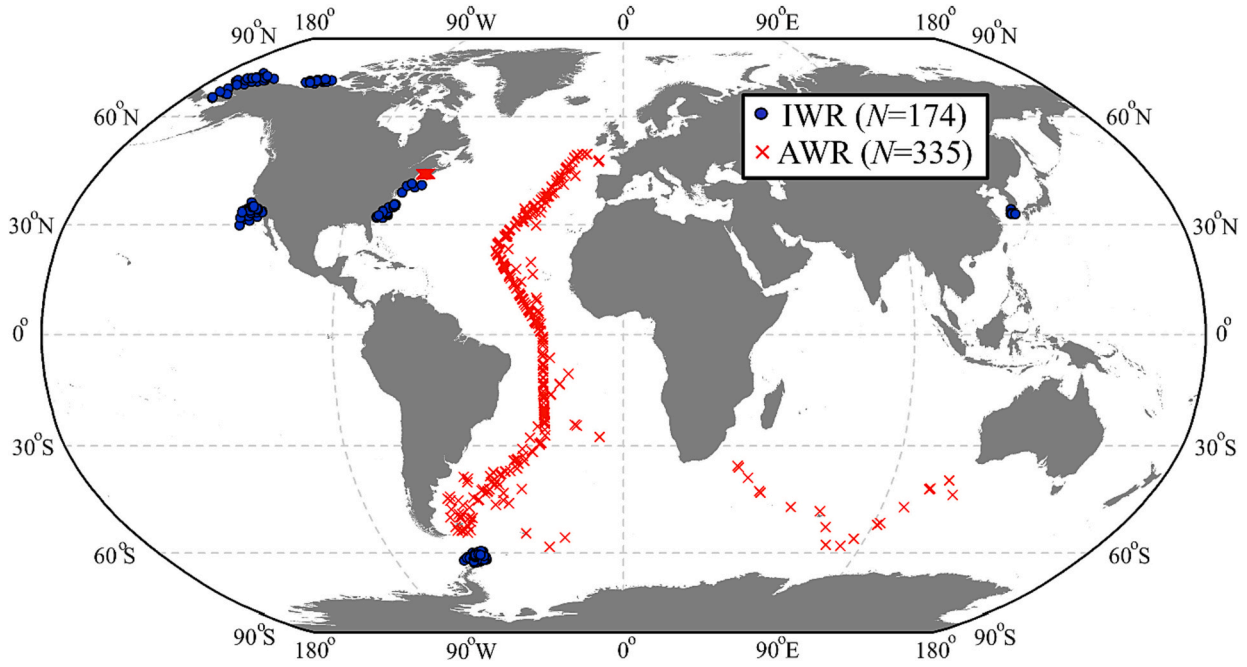

Fig. 1. Geographic location of oceanographic stations in major oceanic basins where coincident field data of spectral remote-sensing reflectance, $R_{rs}(\lambda)$, and surface concentration of particulate organic carbon, POC, were collected for the in situ validation dataset. Two subsets of the in situ validation dataset, in-water radiometry dataset (IWR; $N = 174$) and above-water radiometry dataset (AWR; $N = 335$), which differ by the method of determination of water-leaving radiance are illustrated in blue (circles) and red (crosses) colors, respectively.

2.1.2. Satellite validation datasets

In contrast to recent studies in which large satellite-in situ matchup datasets were assembled using essentially indiscriminate selection of in situ measurements of POC from many cruises and database sources in order to maximize the number of matchup observations (Świrgoń and Stramska, 2015; Evers-King et al., 2017; Liu et al., 2021), we have taken an approach to create the satellite validation datasets by including only the cruises for which in situ POC data were accompanied with fairly complete information on the methodology of POC determinations, especially regarding the acidification treatment and blank correction. We searched for this information in databases and project-specific websites, metadata files, and cruise-specific scientific papers. As a result of this approach, several cruises on which POC measurements were collected and submitted to public databases are not included in our satellite validation datasets. The accepted in situ measurements of surface POC were subject to a satellite-in situ matchup screening process to create satellite sensor-specific validation datasets (SAT_{sensor}) for SeaWiFS ($SAT_{SeaWiFS}$), MODIS-Aqua ($SAT_{MODIS\text{-}A}$), and VIIRS-SNPP ($SAT_{VIIRS\text{-}SNPP}$) ocean color satellite missions.

The matchup screening process for creating the main $SAT_{SeaWiFS}$ dataset is depicted in Fig. 2. We used the SeaWiFS-derived Level-2 Merged Local Area Coverage (MLAC; spatial coverage of 1.1 km at nadir; Reprocessing version 2018.0) product of $R_{rs}(\lambda)$ to obtain the closest (spatially and temporally) satellite matchups with field measurements of surface POC. Initially, we obtained a total of 494 satellite matchups with accepted measurements of POC for the entire SeaWiFS mission (1998–2010). The validity of these initial matchups was further assessed with several matchup screening criteria depicted in Fig. 2. Briefly, the time difference between in situ water collection for POC determination and satellite measurement of $R_{rs}(\lambda)$ was maintained within a $\pm$ 3 h period (Bailey and Werdell, 2006). A 3 $\times$ 3 pixel-box centered at a location of the field station was used as the valid satellite matchup of $R_{rs}(\lambda)$ for a given in situ POC measurement. The pixel-box was further examined for: (i) pixel availability, (ii) outliers (Liu et al., 2004; Pearson et al., 2016), and (iii) spatial homogeneity (Bailey and Werdell, 2006). As indicated in Fig. 2, we also excluded satellite matchups with a solar zenith angle $\geq 70^{o}$, unusual features in the shape of $R_{rs}(\lambda)$ spectra, which may be indicative of gross error (this screening was also applied to MODIS-Aqua but not VIIRS-SNPP because this sensor has less spectral bands in the visible region), and optically-shallow stations.

A total of 356 valid matchups are included in our final $SAT_{SeaWiFS}$ dataset after the application of all matchup-screening criteria. A similar matchup-screening process was applied to MODIS-Aqua (spatial coverage ~1 km at nadir; period 2002–2019) and VIIRS-SNPP (spatial coverage ~750 m at nadir; period 2012–2018) missions, which resulted in the creation of the final $SAT_{MODIS\text{-}A}$ dataset with 343 matchup observations and final $SAT_{VIIRS\text{-}SNPP}$ dataset with 84 matchup observations. Because of the mission overlap, 117 and 60 common stations were also identified in SAT_{sensor} datasets to facilitate a direct comparison of SeaWiFS vs. MODIS-Aqua and MODIS-Aqua vs. VIIRS-SNPP, respectively. The analysis of these satellite validation datasets allows assessment of algorithm performance and characterization of uncertainties of POC estimation from satellite-derived $R_{rs}(\lambda)$ in the actual context of satellite applications of algorithms. This implies that the resulting deviations between the algorithm-derived and measured POC are associated with multiple sources of uncertainties, including satellite radiometric measurements, atmospheric correction, and satellite-in situ matchups, in addition to uncertainties produced by the algorithm formulation itself and in situ determinations of POC. Geographic locations of satellite matchups for SeaWiFS, MODIS-Aqua, and VIIRS-SNPP missions are shown in Fig. 3.

2.1.3. Subsets of satellite validation datasets with additional criteria

In the recent development of satellite sensor-specific hybrid POC algorithms for applications to the global ocean (Stramski et al., 2022), the algorithm development field dataset was assembled using several inclusion/exclusion criteria based on bottom depth, absorption coefficient of chromophoric dissolved organic matter, and POC vs. chlorophyll relationship. These criteria aimed primarily at generating the dataset representative of mean trendlines between surface POC and $R_{rs}(\lambda)$ in the

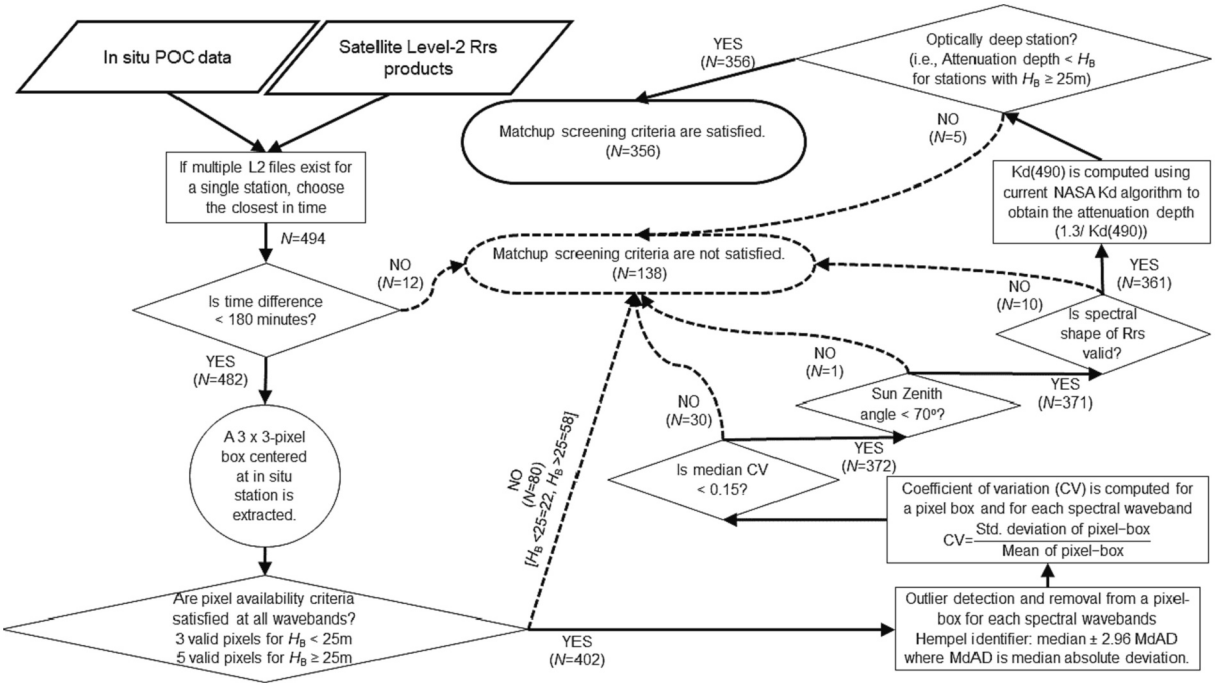

Fig. 2. The matchup screening process to create satellite sensor-specific validation datasets (SAT_{sensor}) for SeaWiFS ($SAT_{SeaWiFS}$). From available 494 satellite-in situ matchups acquired over the SeaWiFS mission, 356 stations satisfied all criteria of the matchup screening process and were assembled to create the $SAT_{SeaWiFS}$ validation dataset. A similar approach was used to search for valid satellite-in situ matchups to create the satellite validation datasets for MODIS-Aqua ($SAT_{MODIS\text{-}A}$; N = 343) and VIIRS-SNPP ($SAT_{VIIRS\text{-}SNPP}$; N = 84) missions.

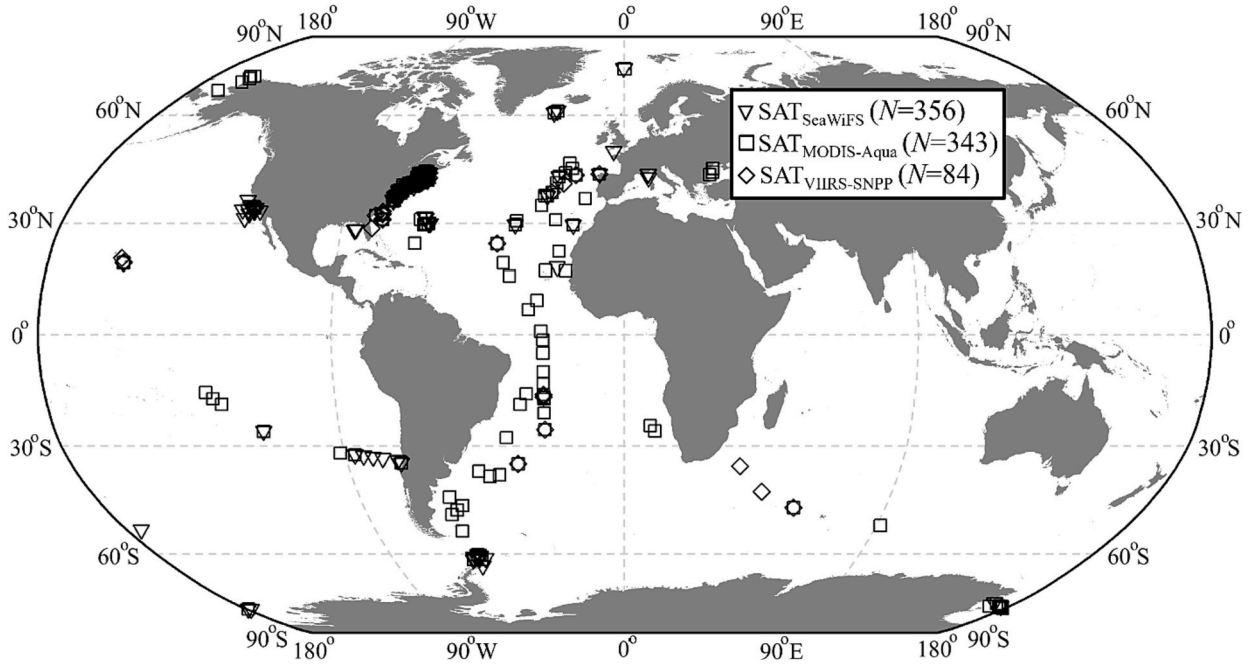

Fig. 3. Geographic locations of satellite-in situ matchups that passed the matchup screening process and were included in the satellite sensor-specific validation datasets (SAT_{sensor}). Three satellite sensor-specific validation datasets are depicted with different symbols; triangles for $SAT_{SeaWiFS}$ ($N = 356$), squares for $SAT_{MODIS\text{-}A}$ ($N = 343$), and diamonds for $SAT_{VIIRS\text{-}SNPP}$ ($N = 84$).

vast open-ocean regions. These bottom depth and bio-optical criteria are described in detail in SJR2022 and are also summarized in Supplementary Material.

For the purposes of this study, the SAT_{sensor} datasets described in Section 2.1.2 were subject to these additional inclusion/exclusion criteria. As a result, we created a subset of $SAT_{SeaWiFS}$ including 310 matchup stations (~13% reduction relative to 356 stations in $SAT_{SeaWiFS}$), a subset of $SAT_{MODIS\text{-}A}$ with 223 stations (~35% reduction relative to 343), and a subset of $SAT_{VIIRS\text{-}SNPP}$ with 49 stations (~42% reduction relative to 84). In summary, the subsets of satellite validation datasets

Table 1
Mathematical formulation for SeaWiFS-specific hybrid POC algorithm (from Stramski et al., 2022) and two other existing global ocean color POC algorithms included for intercomparison analysis. POC range refers to the range of POC values for which the algorithm was formulated.

Algorithm	Mathematical formulation	POC range (mg m^{-3})
Hybrid (SeaWiFS)	If BRDI <1 then $POC = POC_{MBR}$ If BRDI ≥1 then $POC = POC_{MBR}W_{MBR} + POC_{BRDI}W_{BRDI}$ where $W_{MBR} = 0.5[w_{MBR}+(1 - w_{BRDI})]$ and $W_{BRDI} = 1 - W_{MBR}$ BRDI component: $BRDI = \left(\frac{R_{rs}(443) - R_{rs}(555)}{R_{rs}(490)}\right)$ $POC_{BRDI} = 10^{(1.5407+0.8586\ BRDI-0.0787\ BRDI^2-1.8571\ BRDI^3+1.5738\ BRDI^4-0.3839\ BRDI^5)}$ $w_{BRDI} = \left\{\begin{array}{c} 0 \text{ if } POC_{BRDI} > 25 \text{ mg m}^{-3} \\ 1 \text{ if } POC_{BRDI} < 15 \text{ mg m}^{-3} \\ 0 \le 1 - \log_{10}[0.9\ POC_{BRDI} - 12.5] \le 1 \text{ if } 15 \le POC_{BRDI} \le 25 \text{ mg m}^{-3} \end{array}\right\}$ MBR component: $MBR = \log_{10}\left[\text{maximum}\left(\frac{R_{rs}(443)}{R_{rs}(555)}, \frac{R_{rs}(490)}{R_{rs}(555)}, \frac{R_{rs}(510)}{R_{rs}(555)}\right)\right]$ $POC_{MBR} = 10^{(2.5037-2.1297\ MBR+1.8727\ MBR^2-0.9554\ MBR^3)}$ $w_{MBR} = \left\{\begin{array}{c} 1 \text{ if } POC_{MBR} > 25 \text{ mg m}^{-3} \\ 0 \text{ if } POC_{MBR} < 15 \text{ mg m}^{-3} \\ 0 \le \log_{10}[0.9\ POC_{MBR} - 12.5] \le 1 \text{ if } 15 \le POC_{MBR} \le 25 \text{ mg m}^{-3} \end{array}\right\}$	11.9–1022.1
BR-PF[a]	$POC = 203.2\left(\frac{R_{rs}(443)}{R_{rs}(555)}\right)^{-1.034}$	11.9–266.8
CI[b]	$CI = R_{rs}(555) - \left[R_{rs}(490) + \left(\frac{555-490}{670-490}\right)(R_{rs}(670) - R_{rs}(490))\right]$ If $CI \ge 0.0005\ sr^{-1} \rightarrow POC = 10^{(485.19\ CI+2.10)}$ If $CI < 0.0005\ sr^{-1} \rightarrow POC = 10^{(185.75\ CI+1.97)}$	~15–900

[a] BR-PF is the blue-to-green band ratio power function algorithm also referred to as standard POC algorithm (Stramski et al., 2008).
[b] CI is the color-index based POC algorithm (Le et al., 2018).

described in this section are intended to provide additional insights into the performance of global POC algorithms when assessed with validation data that satisfy a certain specific set of criteria consistent with bio-optical properties generally observed in open-ocean pelagic environments and were used in the generation of algorithm development dataset for development of global hybrid POC algorithms in SJR2022.

2.2. Ocean color POC algorithms

While the main focus of this study is to conduct the validation of recently developed global hybrid POC algorithms for applications using different ocean color satellite sensors (Stramski et al., 2022), we also include a comparison of validation results for satellite sensor-specific hybrid algorithms with a few existing global ocean color POC algorithms described in the literature. The formulations of the SeaWiFS-specific hybrid algorithm and other existing POC algorithms used in this comparison are presented in Table 1. Below is a summary of these algorithms.

2.2.1. Global hybrid POC algorithms

The hybrid POC algorithm consists of two formulations, i.e., a fifth-degree polynomial function using the reflectance band ratio difference index (BRDI) applicable to very low POC (< 15 mg m^{-3}) and a third-degree polynomial function using the reflectance maximum band ratio (MBR) applicable to higher POC (> 25 mg m^{-3}) (Table 1, see also Stramski et al., 2022). The weighting approach is applied to these two formulations within a narrow transition range of POC. We note that in addition to SeaWiFS, the hybrid algorithms were also formulated for other ocean color satellite sensors (Stramski et al., 2022). The hybrid algorithms for MODIS and VIIRS-SNPP sensors are also used in the present validation study. In the SeaWiFS-specific hybrid algorithm, the MBR formulation uses a 510 nm band to extend the applicability of the hybrid algorithm in high POC waters; however, MODIS and VIIRS-SNPP sensors do not have this band. A virtual 510 nm band is therefore estimated for MODIS and VIIRS-SNPP sensors to make the MODIS- and VIIRS-specific hybrid algorithms similar to the SeaWiFS-specific hybrid algorithm. R_{rs} at the virtual 510 nm band is obtained from $R_{rs}(488)$ and $R_{rs}(531)$ for the MODIS-specific hybrid algorithm, whereas it is computed from $R_{rs}(486)$ and $R_{rs}(551)$ for the VIIRS-SNPP-specific hybrid algorithm (Stramski et al., 2022).

The global hybrid POC algorithms were formulated with field measurements covering about a 100-fold range of near-surface POC (11.9–1022.1 mg m^{-3}) and representing a variety of water types from five major oceanic basins ranging from ultraoligotrophic waters of the South Pacific Gyre to the productive waters in the Arctic coastal region. The algorithm development dataset was assembled using several criteria to ensure consistency with typical bio-optical properties of vast areas of open-ocean pelagic environments. We also recall that the main suite of hybrid algorithms was developed with POC measurements uncorrected for DOC adsorption as described in SJR2022 and the validation results described in the present study are focused on the use of such POC data.

2.2.2. Other global POC algorithms

First, we considered the Band-Ratio Power Function (BR-PF) algorithm, which uses a power function of the blue-to-green remote-sensing reflectance ratio, $R_{rs}(443)/R_{rs}(555)$, as a predictor of POC (Table 1). This empirical algorithm was obtained from concurrent field measurements of near-surface POC and $R_{rs}(\lambda)$ covering a range of open-ocean pelagic environments from ultraoligotrophic to eutrophic waters of the Pacific and Atlantic Oceans (Stramski et al., 2008). This algorithm has been implemented and is currently used by the NASA Ocean Biology Processing Group to routinely produce global maps of POC from satellite ocean color missions. Thus, this algorithm is also referred to as the standard POC algorithm in this study. Although the original BR-PF algorithm was formulated using SeaWiFS bands, the NASA OBPG also provides a standard global POC product for MODIS-Aqua and VIIRS-SNPP missions. The MODIS and VIIRS-specific versions of the standard POC algorithm have the same coefficients as the SeaWiFS algorithm; however, the NASA Ocean Color Processing Code uses a band-shift approach to shift the band center-wavelength of the available green band in MODIS and VIIRS-SNPP sensors (i.e., 547 nm for MODIS and 551 for VIIRS-SNPP) to the SeaWiFS band of 555 nm (https://oceancolor.gsfc.nasa.gov/atbd/poc/). In this study, in addition to the SeaWiFS algorithms, the standard MODIS- and VIIRS-SNPP-specific BR-PF algorithms were also assessed and compared to the newly developed hybrid algorithms using the in situ and satellite validation datasets.

Second, we considered a Color Index empirical algorithm (referred to as CI algorithm) that was formulated using a global matchup dataset of in situ POC determinations and satellite (SeaWiFS)-derived $R_{rs}(\lambda)$ (Le et al., 2018). The underlying concept of the CI algorithm is similar to the chlorophyll band-difference algorithm (Hu et al., 2012); however, the POC algorithm uses a different combination of three spectral bands (i.e., 490 nm, 555 nm, and 670 nm) in the determination of color index (Table 1). The CI-approach was postulated to be less sensitive to errors in the satellite-derived $R_{rs}(\lambda)$, which can result in improvement of image quality over traditional band-ratio algorithms (Le et al., 2018).

2.3. Statistical metrics for evaluation and intercomparison of different algorithms

To evaluate and compare the performance of the hybrid and two other global POC algorithms, we used a suite of statistical metrics commonly used in ocean bio-optical and biogeochemical studies for characterizing the systematic and relative differences between algorithm-derived and measured values (Brewin et al., 2015; Seegers et al., 2018). For ordinary (untransformed) data of POC, the median bias (*MdB*) and median ratio of algorithm-derived to observed values (*MdR*) were used as a measure of aggregate bias whereas the median absolute

Table 2
Statistical metrics that are used in validation and intercomparison of different POC algorithms.

Symbol	Description
N	Number of samples
x_i, y_i	Observed "x" and model-predicted (algorithm-derived) "y" value for sample i of N
Model evaluation metrics for log-transformed data	
R	Pearson's correlation coefficient between $\log(y_i)$ and $\log(x_i)$
S and I	Slope and intercept of Model II linear regression of $\log(y_i)$ on $\log(x_i)$
MdSA	Median symmetric accuracy (in percent) calculated as $(10^{[\text{median}(\lvert\log(yi)-\log(xi)\rvert]} - 1) \times 100$
MB_{log}	Mean bias calculated as $10^{[\sum(\log(yi)-\log(xi))/N]}$
Model evaluation metrics for untransformed data	
MdR	Median ratio of (y_i / x_i)
MdB	Median bias; median value of $(y_i - x_i)$
MdAPD	Median absolute percentage difference, median value of $100 \times [\lvert(y_i - x_i)/ x_i\rvert]$
RMSD	Root mean square deviation calculated as $[(1/N) \sum_{i=1}^{N}(y_i - x_i)^2]^{0.5}$
% wins	Percentage wins in pairwise comparisons of closeness of x_i and y_i data for a given pair of compared algorithms

percentage difference (*MdAPD*) was used as a measure of relative difference between the algorithm-derived and measured values (Table 2). The root mean square deviation (*RMSD*) is a widely used statistical metric and is also here reported. Two metrics, mean bias (MB_{log}) and median symmetric accuracy (*MdSA*), were used to evaluate the bias and relative difference between log-transformed algorithm-derived POC and log-transformed measured POC (Table 2) (Morley et al., 2016; Morley et al., 2018). These "logarithmic" metrics are unitless and can be interpreted as a percentage difference. Note that for MB_{log} the interpretation of bias in percent is achieved with the expression $(MB_{log} - 1) \times 100$. Importantly, *MdSA* also offers an interpretive advantage over *MdAPD* and *RMSD* because it does not penalize over- and under-prediction differently. The Pearson's correlation coefficient (*R*) and the slope (*S*) and intercept (*I*) of Model-II linear regression (the reduced major axis method) were also computed from the log-transformed algorithm-derived vs. measured POC data (Table 2). A pair-wise comparison method was also used to compare the performance of different algorithms (Broomell et al., 2011; Seegers et al., 2018) and quantify the percent wins for a given examined hybrid algorithm against another examined (BR-PF or CI) algorithm.

3. Results and discussion

3.1. Validation analysis with in situ dataset

3.1.1. Overview of in situ validation dataset

The geographical distribution of oceanographic stations where the in situ validation data were collected covers a broad range of contrasting oceanic environments, including both the continental shelf and off-shelf environments (Fig. 1). For the stations in the IWR dataset, the median value of bottom depth H_B is 678 m and the range is 19–5223 m, which includes 6 stations with $H_B < 25$ m. For the AWR dataset, the median of H_B is 4314 m and the range is 36–6414 m. About 66% of stations in the IWR dataset and 84% of stations in the AWR dataset are in the oceanic off-shelf pelagic zone. The probability distributions of POC for the combined in situ validation dataset and its IWR and AWR subsets are shown in Fig. 4a. In comparison to the probability density function of global satellite-derived POC, which exhibits a maximum for POC between about 20 and 40 mg m^{-3}, the probability distribution of the IWR data shows a shift with the maximum near 70 mg m^{-3} (Fig. 4a). This indicates that the IWR dataset is not consistent with the global POC distribution dominated by conditions of the open-ocean pelagic environments (see also Table S1, Supplementary Material). In contrast, the probability distribution of the AWR dataset exhibits a maximum between 20 and 40 mg m^{-3}, which is consistent with the main maximum of the global distribution derived from SeaWiFS observations (Fig. 4a). However, in comparison to the global SeaWiFS-derived distribution, which also exhibits a secondary peak at about 50 mg m^{-3}, the AWR dataset appears to have insufficient representation of waters where POC values are between about 50 and 100 mg m^{-3} (e.g., the equatorial Pacific and Atlantic Oceans and large portions of the Southern Ocean). Furthermore, the majority of AWR stations are geographically restricted to open-ocean waters in the Atlantic (Fig. 1). The combination of IWR and AWR data provides the in situ validation dataset with an improved representation of POC variability across diverse shelf and open-ocean pelagic environments (Fig. 1), although the probability distribution of this combined dataset (black solid line in Fig. 4a) still shows some notable differences compared to the SeaWiFS-derived distribution.

3.1.2. Validation of SeaWiFS-specific global hybrid POC algorithm and comparison with other POC algorithms

Fig. 5 shows scatter plots of algorithm-derived POC versus measured POC for the SeaWiFS-specific hybrid POC algorithm and two other global algorithms (as described in section 2.2) that are applied to the combined in situ validation dataset. The performance statistics of these algorithms are summarized in Table 3.

As shown in Fig. 5 and Table 3, the hybrid and BR-PF algorithms show generally better performance compared to the CI algorithm. In particular, these two algorithms are characterized by better Model-II regression parameters (*S* and *I*) of the relationship between the algorithm-derived and measured POC (Fig. 5a, b) and improved measures of bias and relative difference (MB_{log}, *MdR*, *MdB*, *MdSA*, and *MdAPD*) in comparison to the CI algorithm (Table 3). Importantly, however, if the hybrid and BR-PF algorithms are directly compared, it is

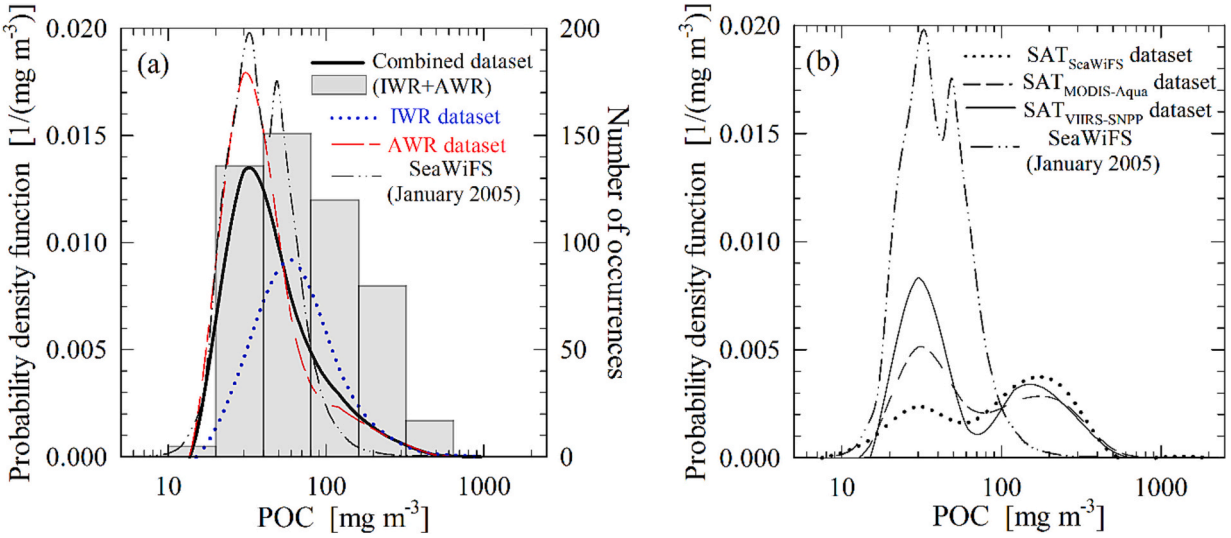

Fig. 4. (a) Probability distributions of POC values obtained from field measurements that are included in the in situ validation dataset. The probability density functions (left y-axis) are shown for the combined in situ validation dataset (i.e., IWR + AWR, number of data $N = 509$) and two subsets of data, IWR ($N = 174$) and AWR ($N = 335$). The histogram (right y-axis) for the combined dataset is also depicted. (b) The probability density functions of POC from field measurements that are included in the satellite-in situ matchup validation datasets for the SeaWiFS, MODIS-Aqua, and VIIRS-SNPP ocean color observations. The satellite validation datasets for the observations with the three satellite ocean color sensors are denoted as $SAT_{SeaWiFS}$ ($N = 356$), $SAT_{MODIS\text{-}A}$ ($N = 343$), and $SAT_{VIIRS\text{-}SNPP}$ ($N = 84$). For comparison, an example probability density function of POC derived from global satellite observations is also shown in both panels. This satellite-derived distribution of POC represents a monthly composite for January 2005 obtained from the application of our SeaWiFS-specific hybrid algorithm (see Section 2.2.1 and Stramski et al., 2022) to SeaWiFS observations.

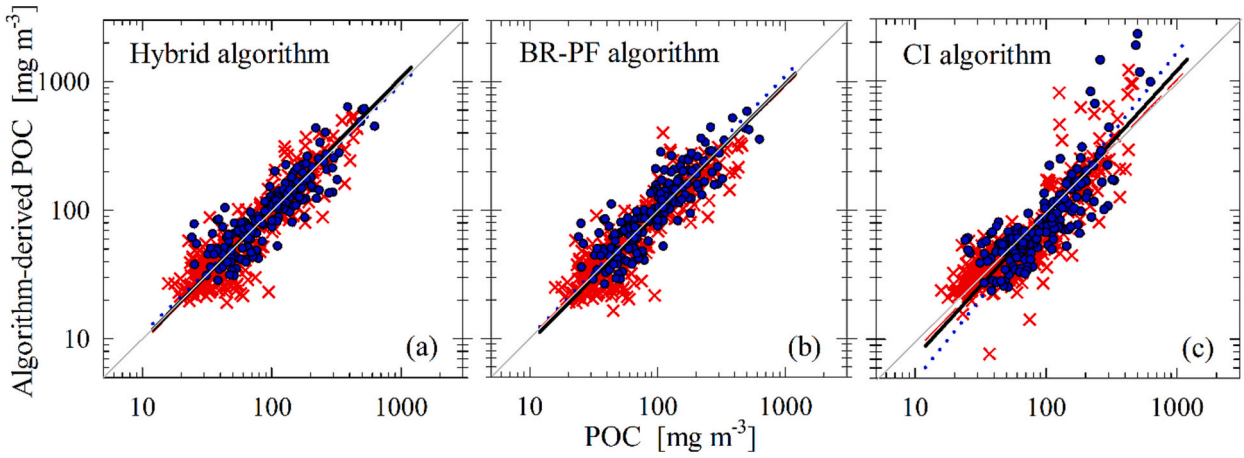

Fig. 5. Algorithm-derived POC vs. measured POC for the in situ validation dataset to assess and intercompare the predictive capabilities of SeaWiFS-specific (a) hybrid algorithm, (b) BR-PF (Band Ratio Power Function) algorithm, (c) CI (Color-Index) algorithm. The two subsets of the combined in situ validation dataset, IWR and AWR, are depicted in blue (circles) and red (crosses) symbols, respectively. Model-II regression lines fitted to the log-transformed data of combined in situ validation dataset ($N = 509$) and the two subsets, IWR ($N = 174$) and AWR ($N = 335$), are shown as a black solid line, a blue dotted line, and a red dashed line, respectively (see Table 3 and Table S2 for regression coefficients and evaluation metrics). The 1:1 line is also depicted (solid grey line).

Table 3
Statistical evaluation parameters and selected results of pair-wise comparisons for satellite sensor-specific global hybrid, BR-PF, and CI algorithms for estimating POC from ocean color. The results were obtained by evaluation of the algorithms with the entire combined in situ validation dataset and its subset restricted to POC > 200 mg m^{-3} (see also Figs. 5 and 6). The statistical parameters are explained in Table 2. % wins are reported for the hybrid algorithm against the other algorithms.

Algorithm	Sensor	Log-transformed data					Untransformed data				% wins
		R	*S*	*I*	*MdSA* (%)	MB_{log}	*MdR* ≡ MdB_{log}	*MdB* (mg m^{-3})	*MdAPD* (%)	*RMSD* (mg m^{-3})	
Results for the entire combined dataset: IWR + AWR (N = 509)											
Hybrid	SeaWiFS	0.91	1.02	−0.03	19.7	1.02	1.00	0.9	18.3	42.8	–
BR-PF	SeaWiFS	0.90	1.01	−0.04	23.0	0.97	0.98	−1.1	20.1	46.7	60
CI	SeaWiFS	0.85	1.11	−0.25	27.2	0.90	0.90	−6.7	22.2	235.5	59
Hybrid	MODIS	0.91	1.05	−0.08	21.0	1.01	1.02	0.8	18.7	46.6	–
BR-PF	MODIS	0.90	0.99	0.02	22.5	1.00	1.00	−0.1	19.5	47.6	58
Hybrid	VIIRS	0.90	1.04	−0.07	20.7	1.02	1.02	1.1	19.0	49.1	–
BR-PF	VIIRS	0.90	1.01	−0.02	23.0	0.97	0.97	−1.5	20.2	47.1	59
Results for a subset of combined dataset with POC > 200 mg m^{-3} (N = 62)											
Hybrid	SeaWiFS	0.64	1.55	1.30	20.4	0.92	0.95	−12.6	18.4	92.7	–
BR-PF	SeaWiFS	0.58	1.30	−0.82	28.5	0.81	0.80	−48.6	22.8	98.0	63
Hybrid	MODIS	0.60	1.49	−1.23	18.7	0.94	0.98	−6.6	18.0	88.6	–
BR-PF	MODIS	0.56	1.26	−0.73	34.1	0.79	0.77	−58.9	26.9	101.3	63
Hybrid	VIIRS	0.59	1.49	−1.22	19.8	0.95	0.98	−5.5	16.7	89.9	–
BR-PF	VIIRS	0.57	1.29	−0.80	31.9	0.79	0.78	−57.9	26.7	100.6	61

clear that the hybrid algorithm has somewhat improved performance. This is an important result, especially in the context of the current use of BR-PF as the standard POC algorithm for the routine processing of ocean color satellite data. Specifically, our validation analysis with the independent in situ dataset shows that the hybrid algorithm yields the POC estimates with essentially no aggregate bias (*MdR* = 1.0) and provides the lowest values of *MdSA* (19.7%), *MdAPD* (18.3%), and *RMSD* (42.8 mg m^{-3}) (Table 3). The pair-wise comparison further supports the best performance of the hybrid algorithm with about 60% wins in the pair-wise comparisons with two other algorithms (Table 3).

It is notable that although the CI algorithm was developed with matchup data of satellite-derived $R_{rs}(\lambda)$ and in situ POC with the intended global applicability, its performance statistics are clearly inferior compared with the hybrid and BR-PF algorithms. In particular, the CI algorithm exhibits a non-linear pattern of algorithm-derived vs. measured POC indicating systematic deviations of POC estimates from the 1:1 line, which are also dependent on the magnitude of POC (Fig. 5c). The aggregate bias is negative (about 10% with *MdR* = 0.9) and *RMSD* is the highest (235.5 mg m^{-3}) among the tested algorithms (Table 3).

Similar to the validation analysis of the combined in situ dataset, we calculated the performance statistics for the IWR and AWR subsets of data separately. The conclusions from these calculations are essentially the same as for the combined dataset. Overall, the global hybrid algorithm has the best statistics for the IWR and AWR datasets, for example, at least 64% wins for the IWR dataset and at least 59% wins for the AWR dataset in pair-wise comparisons against the other two algorithms (Table S2, Supplementary Material).

3.1.3. Validation of MODIS- and VIIRS-SNPP-specific global hybrid POC algorithms and comparison with the standard BR-PF algorithms

Fig. 6 shows scatter plots of algorithm-derived POC versus measured POC for the MODIS- and VIIRS-specific global hybrid and standard BR-PF algorithms that are applied to the combined in situ validation dataset. The performance statistics for these algorithms are given in Table 3.

As shown in Fig. 6 and Table 3, MODIS- and VIIRS-specific hybrid algorithms have performance statistics similar to the BR-PF algorithms. For example, *MdR* for these four algorithms is approximately 1 to within ±0.03 and *MdAPD* is in the range of about 18–20%. Importantly, however, the MODIS- and VIIRS-specific hybrid algorithms produced 58% and 59% wins in the pair-wise comparisons against the respective sensor-specific BR-PF algorithms that are currently used to generate the standard global POC data product from the MODIS-Aqua and VIIRS-SNPP missions by NASA (Table 3). It is also important to note that the

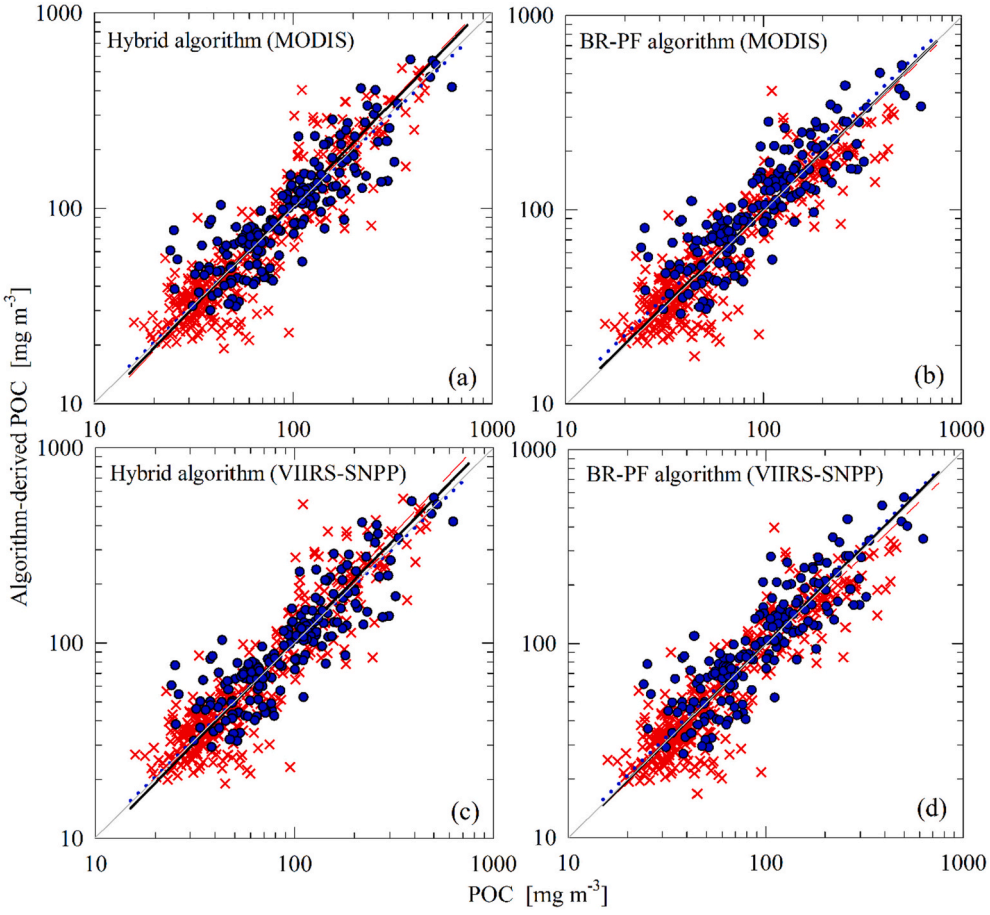

Fig. 6. Algorithm-derived POC vs. measured POC for the in situ validation dataset to assess and intercompare the predictive capabilities of MODIS-specific (a) hybrid algorithm and (b) BR-PF (Band Ratio Power Function) algorithm, and VIIRS-SNPP-specific (c) hybrid algorithm and (d) BR-PF algorithm. The two subsets of the combined in situ validation dataset (i.e., IWR and AWR) are depicted with blue (circles) and red (crosses) colors, respectively. Model-II regression lines for the log-transformed data of combined in situ validation dataset (N = 509) and the two subsets, IWR (N = 174) and AWR (N = 355), are shown as a black solid line, a blue dotted line, and a red dashed line, respectively (see Table 3 and Table S2 in Supplementary Material for regression coefficients and evaluation metrics). The 1:1 line is also depicted (solid grey line).

analysis of the algorithm development dataset used in the SJR2022 study to develop the global hybrid algorithms showed that the hybrid and BR-PF algorithms yield very similar estimates of POC within the range of intermediate POC. This intermediate range is approximately 50–200 mg m^{-3} or even broader 25–200 mg m^{-3} depending on which sensor-specific data are considered (see Fig. 10b for SeaWiFS-specific data in Stramski et al., 2022). In the combined in situ validation dataset, most POC measurements (84% of stations) fall into the POC range of 25–200 mg m^{-3}, which indicates that similar performance statistics for the hybrid and BR-PF algorithms for this validation dataset is consistent with the expectation based on the analysis of the algorithm development dataset in SJR2022.

The combined in situ validation dataset contains a very small number of measurements with very low POC < 25 mg m^{-3} (~4% of stations) and a relatively small number of measurements (~12% of stations) with relatively high POC > 200 mg m^{-3}, which represent the "low" and "high" POC ranges where the hybrid algorithms are expected to provide the most significant improvements in POC estimates compared to the BR-PF algorithms (Stramski et al., 2022). It is thus instructive to compare the performance of hybrid and BR-PF algorithms for the subset of combined in situ validation dataset containing stations with POC > 200 mg m^{-3} (N = 62). The results of this comparison are included in Table 3 for SeaWiFS-, MODIS-, and VIIRS-SNPP-specific algorithms and they clearly support the superiority of hybrid algorithms within the range of relatively high POC. In particular, the satellite sensor-specific hybrid algorithms are characterized by a small aggregate bias with *MdR* values of 0.95–0.98 compared to significant negative bias for sensor-specific BR-PF algorithms with *MdR* of 0.77–0.80. The hybrid algorithms also have improved statistics for the relative difference (*MdSA* and *MdAPD*) and *RMSD*; for example, *MdSA* is 19–20% compared to 28–34% for the BR-PF algorithms (Table 3). Finally, the pair-wise comparisons further support the superiority of hybrid algorithms (61–63% wins) over BR-PF algorithms.

The improvements in the performance of hybrid algorithms compared to BR-PF algorithms at relatively high POC can be attributed to two main distinctive differences between these algorithms. First, these algorithms are based on different mathematical formulations that link POC to the predictor variable involving $R_{rs}(\lambda)$, and second, they use different predictor variables (Stramski et al., 2022). Specifically, the BR-PF uses a single blue-to-green band ratio of $R_{rs}(\lambda)$ under all conditions regardless of POC, for example, $R_{rs}(443)/R_{rs}(555)$ for SeaWiFS. In contrast, the hybrid algorithms use the maximum band ratio (MBR) selected from three band ratios over a broad range of POC > 25 mg m^{-3}, and this component of the hybrid algorithms is referred to as MBR-OC4 for SeaWiFS and MBR-OC4v for MODIS and VIIRS sensors (Stramski et al., 2022). The important point, especially in the context of POC estimates at relatively high POC, is that one of the three band ratios involved in the MBR-OC4 or MBR-OC4v is the ratio that has $R_{rs}(510)$ in the numerator. For SeaWiFS, it is the actually measured $R_{rs}(510)$. Because the 510-nm band is not available for MODIS and VIIRS, for these sensors $R_{rs}(510v)$ at the virtual 510-nm band is estimated from measurements at neighboring bands available on these sensors (Stramski et al., 2022). When POC > 200 mg m^{-3}, the band ratio with $R_{rs}(510)$ or $R_{rs}(510v)$ in the numerator is often selected as the maximum band ratio in the MBR-OC4 or MBR-OC4v component of hybrid algorithms, which generally yields improved estimates of POC compared to BR-PF estimates. This validation analysis also supports the concept of using the virtual 510-nm band in the MODIS- and VIIRS-specific hybrid algorithms as these algorithms provide significant improvements compared to the respective BR-PF algorithms, similarly to the case of SeaWiFS equipped with the 510-nm band. Another noteworthy result for POC > 200 mg m^{-3} in Table 3 is the greater improvement of hybrid algorithms

over BR-PF algorithms for MODIS and VIIRS compared to SeaWiFS. Our analysis indicated that this is largely related to how the standard BR-PF algorithms were parameterized for these sensors. The original parametrization was developed for SeaWiFS (Stramski et al., 2008) and this parameterization was subsequently adopted by NASA also for applications to MODIS and VIIRS by "shifting" the measured reflectance at the green bands of MODIS (i.e., 547 nm) and VIIRS-SNPP (i.e., 551 nm) to the green band of SeaWiFS (i.e., 555 nm). This shift-based estimation of $R_{rs}(555)$ can potentially degrade the performance of MODIS and VIIRS standard BR-PF algorithms which, in turn, can contribute to the observed greater improvement of hybrid algorithms over the standard algorithms for these two sensors in comparison to SeaWiFS.

Similar to the validation analysis of the combined in situ dataset, we also calculated the performance statistics of MODIS- and VIIRS-SNPP-specific algorithms for the IWR and AWR subsets of data separately. The conclusions from these calculations are essentially the same as for the combined dataset. Overall, the MODIS- and VIIRS-SNPP-specific global hybrid algorithms have better statistics for the IWR and AWR datasets than the respective sensor-specific BR-PF algorithms, for example, 60% wins for the IWR dataset and 56% wins for the AWR dataset in pair-wise comparisons (Table S2, Supplementary Material).

3.2. Validation analysis with satellite-in situ matchup datasets

3.2.1. Overview of satellite validation datasets

Similar to the in situ validation dataset, the geographical distribution of oceanographic stations where field measurements of POC were matched with satellite observations to create the satellite sensor-specific validation datasets covers a broad range of contrasting oceanic environments, including both the continental shelf and off-shelf environments (Fig. 3). The probability distributions of POC for the $SAT_{SeaWiFS}$, $SAT_{MODIS\text{-}A}$, and $SAT_{VIIRS\text{-}SNPP}$ validation datasets are shown in Fig. 4b. The distributions are bimodal. The location of the first maximum between 20 and 40 mg m^{-3} is consistent with the main maximum of the global POC distribution as illustrated by the example result for January 2005 derived from SeaWiFS observations (Fig. 4b). The second maximum in the distributions of satellite validation datasets is located within the range of 150–200 mg m^{-3}. This feature indicates an over-representation of waters with relatively high POC between about 150 and 500 mg m^{-3} compared with the global satellite-derived POC distribution. This over-representation of relatively high POC is accompanied by the under-representation of lower POC in the distributions of the validation datasets relative to the global distribution. Despite these differences in the shape of probability distributions, the sensor-specific validation datasets cover a significant range of POC, which provides a reasonable basis for evaluating the sensor-specific hybrid and other algorithms with satellite-in situ data matchups.

3.2.2. Validation of SeaWiFS-specific global hybrid POC algorithm and comparison with other POC algorithms

Fig. 7 shows scatter plots of algorithm-derived versus measured POC for the SeaWiFS satellite-in situ matchup validation dataset ($SAT_{SeaWiFS}$; $N = 356$). These scatter plots are presented for the SeaWiFS-specific hybrid, BR-PF, and CI algorithms. The comparison of the performance statistics shown in Table 4 indicates that the hybrid algorithm performed better than the two other algorithms. Specifically, the hybrid algorithm has the lowest values of *MdSA* (21.6%), *MdAPD* (19.4%), and *RMSD* (84.7 mg m^{-3}) (Table 4). Importantly, the hybrid algorithm exhibits significantly lower *RMSD* in comparison to the standard BR-PF algorithm (104.9 mg m^{-3}) as well as the CI algorithm (271.0 mg m^{-3}). The increased *RMSD* of the CI algorithm may be largely associated with the use of red band (i.e., 670 nm for SeaWiFS) in the color index. The pair-wise comparison of algorithms further supports the superior performance of the hybrid algorithm. The hybrid algorithm produced 53% and 52% wins against the BR-PF and CI algorithms, respectively (Table 4). In addition, the distribution of data points in scatter plots in Fig. 7 also indicates comparatively good performance of the hybrid algorithm across the whole range of POC. For example, in contrast to the hybrid algorithm, the standard BR-PF algorithm appears to exhibit a more pronounced tendency to overpredict low values of POC and underpredict the high values.

The scatter plots in Fig. 7 provide important additional insights into the validation analysis. Specifically, the data points shown as green (filled) symbols represent the subset of the $SAT_{SeaWiFS}$ validation dataset that satisfies a specific set of inclusion criteria based on bottom depth ($H_B \geq 50$ m) and bio-optical properties of seawater generally observed in open-ocean pelagic environments (see Section 2.1.3 and Supplementary Material for details). The most notable effect of the application of these criteria to the $SAT_{SeaWiFS}$ dataset is that the data points satisfying these criteria exhibit reduced scatter around the 1:1 line compared to the entire dataset (Fig. 7). It is clear that a number of data points that do not satisfy these criteria (open triangles in Fig. 7) show the largest deviation

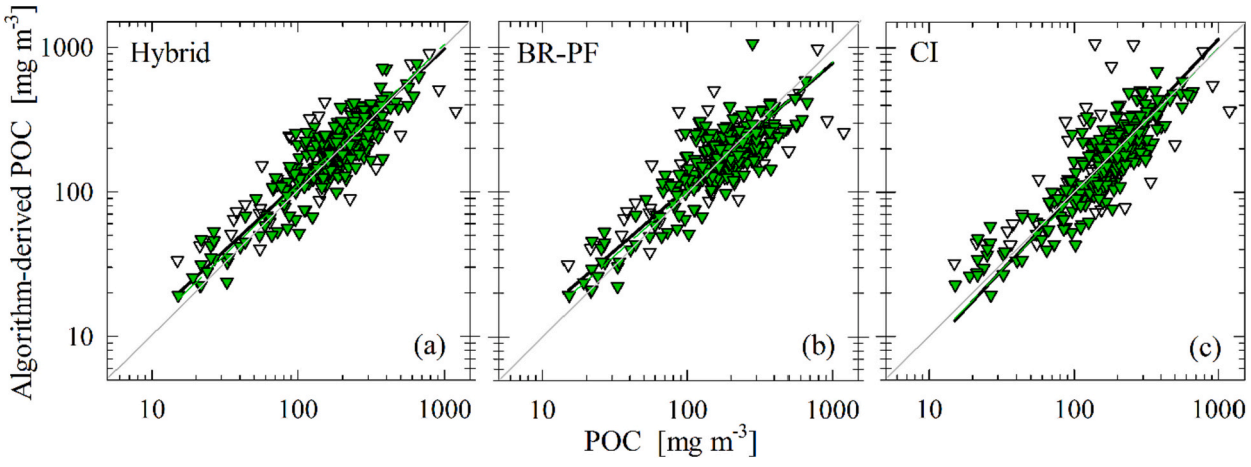

Fig. 7. Algorithm-derived POC vs. measured POC for the SeaWiFS-specific satellite validation dataset ($SAT_{SeaWiFS}$) to assess and intercompare the predictive capabilities of SeaWiFS-specific (a) hybrid algorithm, (b) BR-PF (Band Ratio Power Function) algorithm, (c) CI (Color-Index) algorithm. The data for the entire $SAT_{SeaWiFS}$ dataset ($N = 356$) are depicted by both the open symbols and the green (filled) symbols. The green (filled) symbols represent a subset of $SAT_{SeaWiFS}$ dataset ($N = 310$) that satisfied additional inclusion criteria associated with bottom depth and seawater bio-optical properties generally observed in open-ocean pelagic environments (see Section 2.1.3 and Supplementary Material for details). Model-II regression lines for the log-transformed data of the entire $SAT_{SeaWiFS}$ dataset (solid black lines), the subset satisfying additional inclusion criteria (green dashed lines), and the 1:1 line (grey solid lines) are shown (see Table 4 and Table S3 in Supplementary Material for regression coefficients and evaluation metrics). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Statistical evaluation parameters and selected results of pair-wise comparisons for satellite sensor-specific global hybrid, BR-PF, and CI algorithms for estimating POC from ocean color. The results were obtained by evaluation of the algorithms with the satellite-in situ matchup datasets. The statistical parameters are explained in Table 2. % wins are reported for the hybrid algorithm against the other algorithms.

Algorithm	Sensor	Log-transformed data					Untransformed data				% wins
		R	*S*	*I*	*MdSA* (%)	MB_{log}	*MdR* ≡ MdB_{log}	*MdB* (mg m^{-3})	*MdAPD* (%)	*RMSD* (mg m^{-3})	
Results for the $SAT_{SeaWiFS}$ dataset (N = 356)											
Hybrid	SeaWiFS	0.88	0.93	0.20	21.6	1.10	1.09	14.6	19.4	84.7	–
BR-PF	SeaWiFS	0.83	0.86	0.31	24.9	0.99	0.97	−2.8	21.3	104.9	53
CI	SeaWiFS	0.82	1.07	−0.15	24.2	1.02	0.99	−1.2	21.1	271.0	52
Results for the $SAT_{MODIS-A}$ dataset (N = 343)											
Hybrid	MODIS	0.84	1.03	−0.05	30.8	1.06	1.03	2.7	26.0	146.6	–
BR-PF	MODIS	0.77	0.93	0.13	33.6	0.93	0.87	−15.6	26.9	204.7	59
Results for the $SAT_{VIIRS-SNPP}$ dataset (N = 84)											
Hybrid	VIIRS	0.88	0.92	0.19	33.8	1.04	1.01	0.0	27.4	73.9	–
BR-PF	VIIRS	0.88	0.82	0.34	39.0	0.92	0.85	−21.5	28.4	76.0	64

from the 1:1 line. This result is particularly important from the standpoint of a global hybrid algorithm because, as might have been expected, it provides direct evidence for improvements in performance statistics when the global algorithm is validated with satellite-in situ data matchups that are consistent with the algorithm development data in terms of representativeness of the range of environmental conditions. For example, the analysis of this subset of validation data for the hybrid algorithm in Fig. 7a yields improved values of *MdB* (10.9 mg m^{-3}), *MdSA* (18.4%), *MdAPD* (17.7%), and *RMSD* (63.4 mg m^{-3}) compared to the corresponding values in Table 4, which are based on the entire $SAT_{SeaWiFS}$ dataset. These improvements are substantial, especially considering that only 13% of data in the entire $SAT_{SeaWiFS}$ dataset did not satisfy the bottom depth and bio-optical inclusion criteria. It is notable that the standard global BR-PF algorithm has also improved performance statistics when evaluated with validation data satisfying the inclusion criteria characteristic of open-ocean pelagic environments (i.e., green (filled) symbols in Fig. 7b); for example, *MdSA* was reduced to 22.4% and *RMSD* to 82.4 mg m^{-3} compared to 24.9% and 104.9 mg m^{-3} for the entire $SAT_{SeaWiFS}$ dataset. Table S3 (Supplementary Material) provides statistical parameters for the hybrid, BR-PF, and CI algorithms examined with the subset of SeaWiFS validation data (i.e., green (filled) triangles in Fig. 7). For this subset of data, the improvements in performance statistics are generally seen for all three algorithms.

3.2.3. Validation of MODIS- and VIIRS-SNPP-specific global hybrid POC algorithms and comparison with the standard BR-PF algorithms

Fig. 8 depicts scatter plots of algorithm-derived POC versus measured POC for the MODIS-specific and VIIRS-specific global hybrid

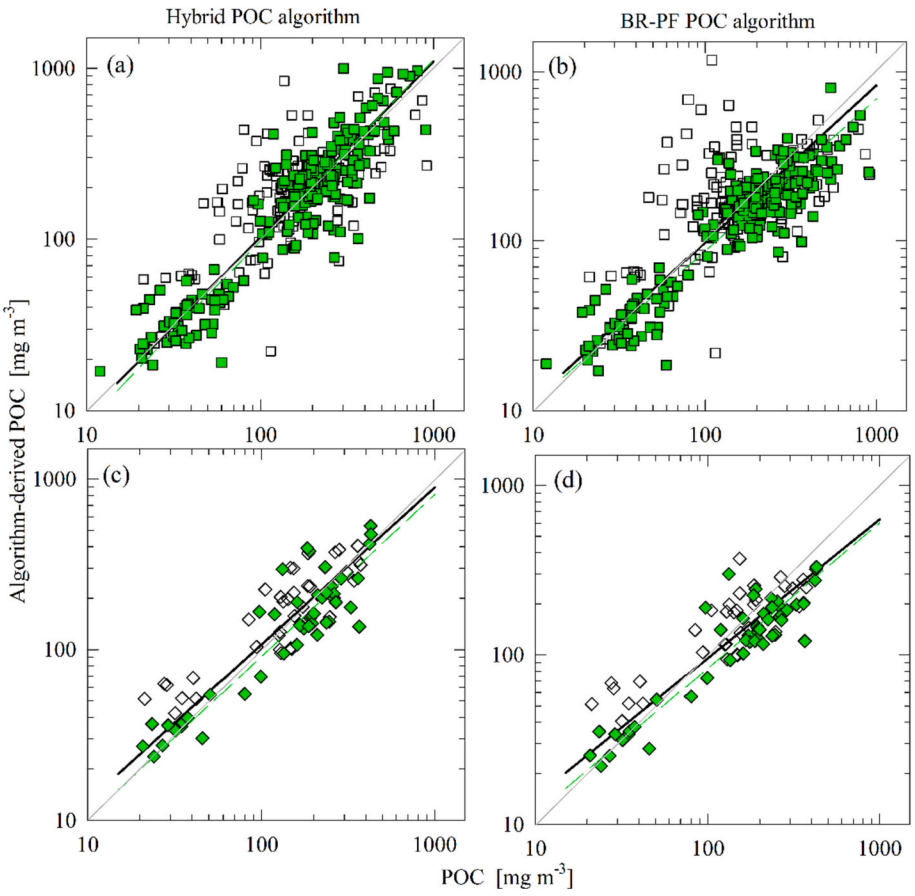

Fig. 8. Algorithm-derived POC vs. measured POC for the sensor-specific satellite validation datasets (SAT_{sensor}) to assess and intercompare the skill of (a) MODIS-specific hybrid algorithm, (b) MODIS-specific BR-PF (Band Ratio Power Function) algorithm, (c) VIIRS-SNPP-specific hybrid algorithm, and (d) VIIRS-SNPP-specific BR-PF algorithm. The data for the entire $SAT_{MODIS-A}$ ($N = 343$) and the entire $SAT_{VIIRS-SNPP}$ ($N = 84$) datasets are depicted by both the open and the green (filled) symbols. The green symbols represent a subset of $SAT_{MODIS-A}$ dataset ($N = 223$) and a subset of $SAT_{VIIRS-SNPP}$ dataset ($N = 49$) that satisfied additional inclusion criteria associated with bottom depth and seawater bio-optical properties generally observed in open-ocean pelagic environments (see Section 2.1.3 for details). Model-II regression lines for the log-transformed data of the entire $SAT_{MODIS-A}$ and $SAT_{VIIRS-SNPP}$ datasets (solid black lines), their subsets satisfying additional inclusion criteria (green dashed lines), the 1:1 line (solid grey lines) are shown (see Table 4 and Table S3 in Supplementary Material for regression coefficients and evaluation metrics). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

algorithms and standard BR-PF algorithms that are applied to the satellite-in situ matchup validation datasets $SAT_{MODIS\text{-}A}$ and $SAT_{VIIRS\text{-}SNPP}$. The performance statistics for these algorithms are given in Table 4. These results for MODIS- and VIIRS-specific algorithms are analogous to those presented in Fig. 7a, b and Table 4 for the SeaWiFS-specific algorithms. Both Fig. 8 and the statistics in Table 4 indicate that the MODIS- and VIIRS-specific hybrid algorithms have better performance in comparison to the BR-PF algorithms. In particular, the hybrid algorithms are characterized by a small aggregate bias with *MdR* values close to 1 compared to significant negative bias for the BR-PF algorithms with *MdR* of 0.85–0.87 (Table 4). The MODIS-specific BR-PF algorithm has also significantly higher *RMSD* than the hybrid algorithm. Moreover, the hybrid algorithms show improved performance at relatively low and high POC in comparison to the BR-PF algorithms as evidenced by data patterns and the associated Model-II regression fits in Fig. 8, as well as the values of *S* and *I* parameters of these regression fits in Table 4. In the pair-wise comparison analysis, the hybrid algorithms recorded 59% and 64% wins over the BR-PF algorithms for the $SAT_{MODIS\text{-}A}$ and $SAT_{VIIRS\text{-}SNPP}$ datasets, respectively.

Similar to Fig. 7, the subsets of the $SAT_{MODIS\text{-}A}$ and $SAT_{VIIRS\text{-}SNPP}$ datasets that satisfy the criteria associated with bottom depth and bio-optical properties typical of open-ocean pelagic environments are identified in Fig. 8 with green (filled) symbols. The validation analyses of these subsets of data yield improved performance statistics for the hybrid and BR-PF algorithms compared to the analyses performed on the entire $SAT_{MODIS\text{-}A}$ and $SAT_{VIIRS\text{-}SNPP}$ datasets. These improvements are especially well pronounced for the MODIS-specific algorithms. For example, the analysis of this subset of validation data for the MODIS-specific hybrid algorithm in Fig. 8a yields improved values of *MdSA* (24.8%), *MdR* (0.99), and *RMSD* (104.1 mg m^{-3}) in comparison to the corresponding values in Table 4, which are based on the entire $SAT_{MODIS\text{-}A}$ dataset. More information on statistical parameters for the MODIS- and VIIRS-specific algorithms examined with the subset of validation data (i.e., green (filled) symbols in Fig. 8) is provided in Table S3 (Supplemental Material).

3.3. Assessment of the consistency between satellite sensor-specific hybrid algorithms

Despite some differences in the center wavelengths of spectral bands of satellite ocean color sensors and the number of available bands that were used to formulate the various satellite sensor-specific hybrid algorithms (Stramski et al., 2022), the algorithm evaluation metrics calculated from the analysis of in situ validation dataset (Table 3 and Table S2 in Supplementary Material) showed the similar skill of SeaWiFS-, MODIS-, and VIIRS-SNPP-specific hybrid algorithms for estimating surface POC from measurements of spectral remote-sensing reflectance, $R_{rs}(\lambda)$. In particular, the three hybrid algorithms provided, on average, nearly unbiased estimates of POC relative to measured POC, and the algorithm-derived POC was obtained with a relative difference of less than about 20% for more than half of the data in the in situ validation dataset (Table 3). The observed differences between the sensor-specific algorithms could have resulted, in part, from measurement and data processing errors in $R_{rs}(\lambda)$ at one or more wavelengths that are used in the sensor-specific hybrid algorithms. Furthermore, the additional uncertainties associated with the use of the virtual 510 nm band in the MODIS- and VIIRS-specific hybrid algorithms could have potentially contributed to the observed differences between the satellite sensor-specific hybrid algorithms. We note that *MdAPD* for the SeaWiFS-specific hybrid algorithm was 13.6% when the algorithm was evaluated using the same in-water radiometry dataset that was used to formulate the algorithm (Stramski et al., 2022). In this study, the application of the SeaWiFS-specific hybrid algorithm to the independent in situ validation dataset (combining both in-water and above-water radiometry) resulted in *MdAPD* of 18.3%, and when this algorithm was evaluated with only in-water radiometry (IWR) data, *MdAPD* was about 22% (Table 3 and Table S2 in Supplementary Material).

When assessed with the satellite-in situ data matchups (i.e., SAT_{sensor} validation datasets), the skill of the satellite sensor-specific hybrid algorithms showed notable differences in algorithm evaluation metrics. For example, whereas the SeaWiFS-specific hybrid algorithm showed relatively better metrics of relative difference between algorithm-derived and measured POC, the MODIS-specific and VIIRS-specific hybrid algorithms were characterized by comparatively better aggregate bias metrics (Table 4). The application of SeaWiFS-, MODIS-, and VIIRS-specific hybrid algorithms to corresponding satellite validation datasets resulted in *MdAPD* of about 19%, 26%, and 27%, respectively (Table 4). Furthermore, a comparison of the values of the correlation coefficient, *R*, and the Model-II regression parameters, *S* and *I*, indicates inferior relationships between sensor-specific algorithm-derived and measured POC for the satellite validation datasets in comparison to the in situ validation dataset (Tables 3 and 4).

Additional insights into the inter-sensor comparison of the performance of satellite sensor-specific hybrid algorithms is provided in Fig. 9. Fig. 9a compares POC derived from the SeaWiFS- and MODIS-specific hybrid algorithms using input data of $R_{rs}(\lambda)$ from the in situ validation dataset. Fig. 9b shows similar results but for comparison of POC estimates from the VIIRS- and MODIS-specific algorithms. Both comparisons show a very good agreement of POC estimates obtained from the sensor-specific algorithms. Specifically, the regression fits to the compared data are very close to the 1:1 line. The slope *S* and intercept *I* of the regression fits are very close to 1 and 0, respectively, and the correlation coefficient *R* is nearly 1. The closeness of agreement between the POC estimates obtained from the different satellite sensor-specific algorithms is also supported by other statistical parameters shown in Fig. 9a,b. The *MdAPD* of about 1–2% indicates that the differences between the compared POC estimates are, on average, very small, and the *MdR* and *MdB* values indicate that the differences exhibit essentially no bias.

Fig. 9c,d also depicts similar comparisons of POC estimates obtained from the different sensor-specific algorithms, but these results were obtained using satellite-derived $R_{rs}(\lambda)$ as input to the algorithms. For comparison of the SeaWiFS- and MODIS-derived POC (Fig. 9c) we used 117 input values of SeaWiFS- and MODIS-derived $R_{rs}(\lambda)$ that represent the common matchups from the $SAT_{SeaWiFS}$ and $SAT_{MODIS\text{-}A}$ datasets corresponding to the same in situ POC value. Similarly, for comparison of the VIIRS- and MODIS-derived POC (Fig. 9d), 60 common measurements of $R_{rs}(\lambda)$ from the $SAT_{VIIRS\text{-}SNPP}$ and $SAT_{MODIS\text{-}A}$ datasets were used as input to the sensor-specific algorithms. The data in Fig. 9c,d indicate that the inter-sensor agreement between POC retrievals from satellite measurements of $R_{rs}(\lambda)$ is not as good as the agreement between POC estimates based on in situ measurements of $R_{rs}(\lambda)$ shown in Fig. 9a,b. This result is not surprising because satellite-derived $R_{rs}(\lambda)$ is subject to various additional sources of uncertainty, most notably the atmospheric correction, which are irrelevant to in situ determinations of $R_{rs}(\lambda)$ (Hu et al., 2001; Hu et al., 2013; Frouin et al., 2019; Joshi and D'Sa, 2020). Naturally, the satellite inter-sensor comparisons are also affected by inherent differences in satellite sensor specifications such as radiometric and spectral characteristics. In addition, the pairs of $R_{rs}(\lambda)$ values derived from two satellite sensors and used for inter-sensor comparisons of POC retrievals are not matched perfectly in space and time, which may also have some effect on the scatter of data in Fig. 9c,d (see also Pahlevan et al., 2016). Notwithstanding the fact that multiple factors other than the formulation of satellite sensor-specific algorithms themselves have an effect on results in Fig. 9c,d, these inter-sensor comparisons are encouraging in terms of demonstrating reasonably good agreement between satellite retrievals of POC from different satellite ocean color missions. Specifically, on average, the differences between the sensor-specific POC retrievals remain relatively low, below 20%, with *MdAPD* of about 10.8% for the differences between the SeaWiFS- and MODIS-specific retrievals (Fig. 9c), and 16.8% between the VIIRS- and MODIS-specific retrievals (Fig. 9d). It is also important to note that

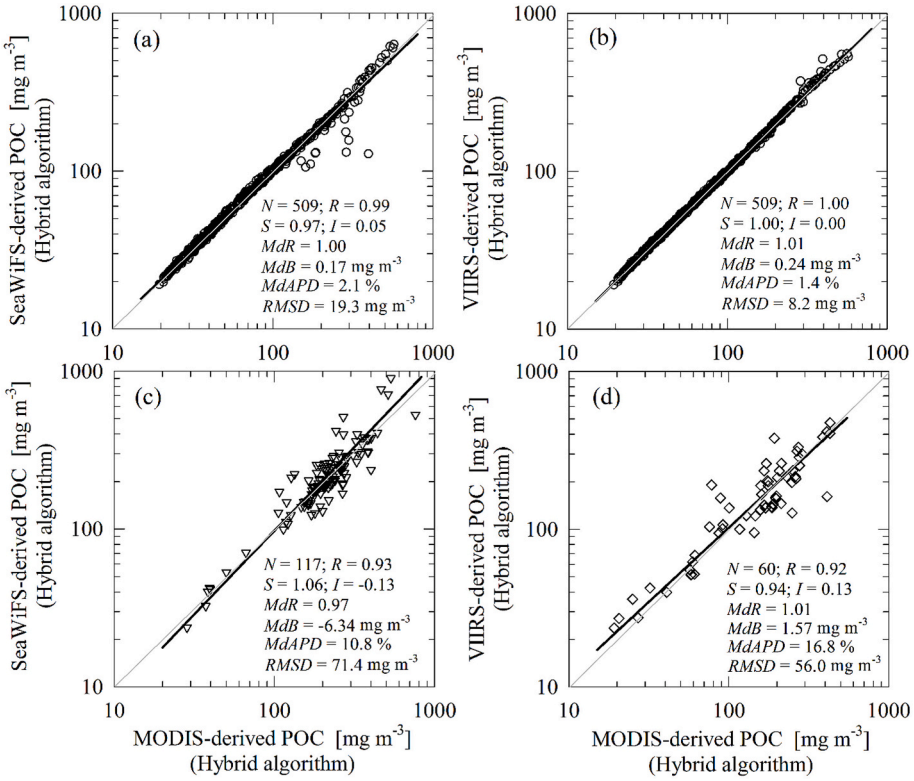

Fig. 9. Inter-sensor comparison of algorithm-derived POC from satellite sensor-specific algorithms. (a) Comparison of SeaWiFS-specific and MODIS-specific hybrid algorithms using input remote-sensing reflectance $R_{rs}(\lambda)$ from the in situ validation dataset. (b) Comparison of VIIRS-SNPP-specific and MODIS-specific hybrid algorithms using input $R_{rs}(\lambda)$ from the in situ validation dataset. (c) Same as panel (a) but using satellite-derived $R_{rs}(\lambda)$ as input to the sensor-specific hybrid algorithms. These input data of $R_{rs}(\lambda)$ represent the common satellite-in situ matchup stations in the satellite validation datasets $SAT_{SeaWiFS}$ and $SAT_{MODIS\text{-}A}$. (d) Same as panel (b) but using satellite-derived $R_{rs}(\lambda)$ as input to the sensor-specific hybrid algorithms. These input data of $R_{rs}(\lambda)$ represent the common satellite-in situ matchup stations in the satellite validation datasets $SAT_{VIIRS\text{-}SNPP}$ and $SAT_{MODIS\text{-}A}$. Model-II regression lines for the log-transformed data (solid black lines) and the 1:1 line (grey lines) are shown along with several statistical parameters that are defined in Table 2.

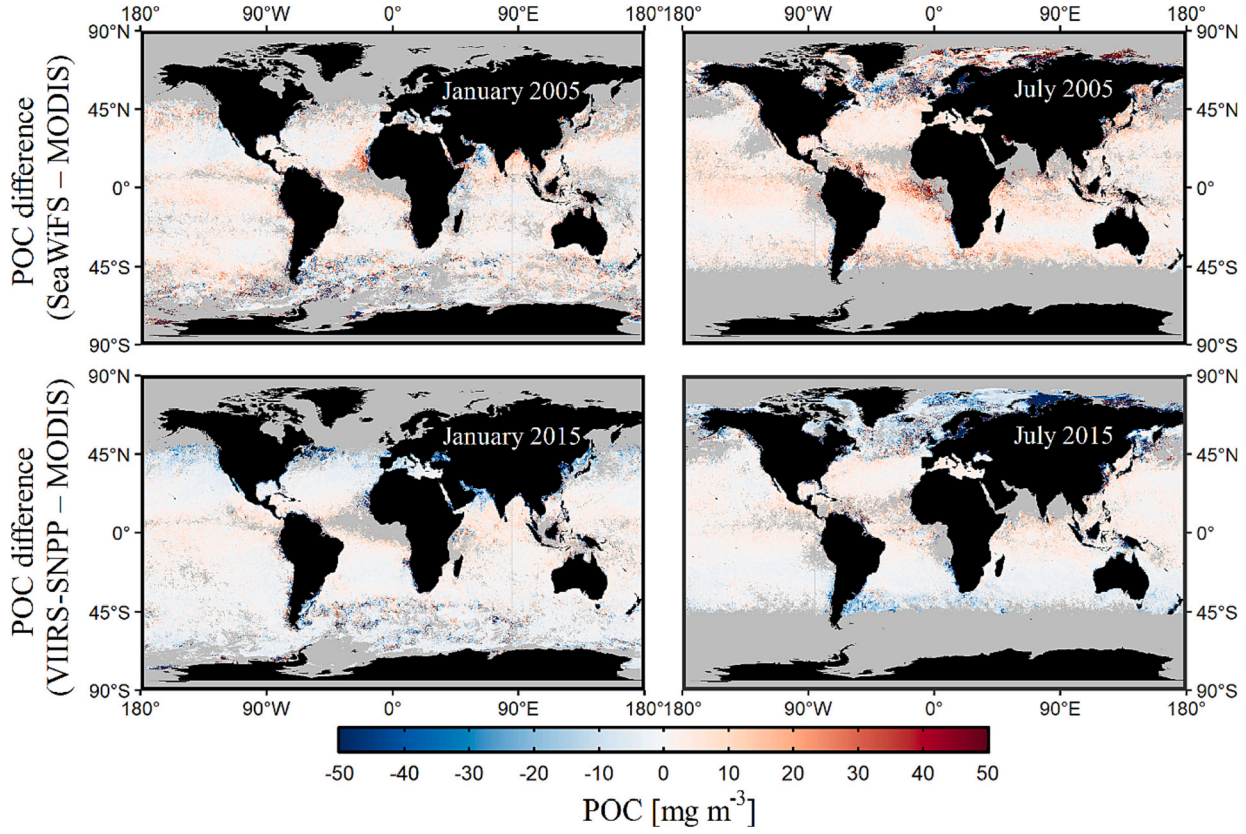

Fig. 10. (Top panels) The difference between POC values derived from the SeaWiFS-specific hybrid algorithm and MODIS-specific hybrid algorithm for two example months, January 2005 and July 2005, as calculated from monthly composites of $R_{rs}(\lambda)$ data product for each sensor. (Bottom panels) The difference between POC values derived from the VIIRS-SNPP-specific hybrid algorithm and MODIS-specific hybrid algorithm for January 2015 and July 2015. The $R_{rs}(\lambda)$ data for the MODIS-specific algorithm were obtained from the MODIS-Aqua satellite mission.

the inter-sensor differences in satellite retrievals of POC exhibit a very small or negligible aggregate bias as indicated by *MdR* values very close to 1 in Fig. 9c,d. Although the results in Fig. 9 are indicative of the reasonably high degree of inter-sensor consistency of POC retrievals, the comparisons in Fig. 9c,d are admittedly limited in terms of the number of data included in our satellite-in situ matchup validation datasets.

More extensive inter-sensor comparative analyses with the use of large satellite datasets covering the entire overlapping periods between different satellite missions are desirable in the near future to support the long-term consistency of POC products from multiple satellite missions. Here we illustrate example satellite images of the difference between POC derived from the SeaWiFS- and MODIS-specific hybrid algorithms, and between VIIRS-SNPP- and MODIS-specific hybrid algorithms (Fig. 10). For this illustration, the monthly composite images of satellite-derived $R_{rs}(\lambda)$ were used as input to the sensor-specific POC algorithms for the months of January and July from 2005 and 2015. In general, the differences in POC derived from different satellite sensors are small, especially over vast areas of the open ocean at subtropical and tropical latitudes. Based on these data over the global ocean, the median difference between POC derived from the SeaWiFS- and MODIS-specific algorithms has slightly positive value below 3 mg m^{-3} and *MdAPD* <9% for the two example months in this comparison. These statistical parameters for the VIIRS-SNPP-derived POC and MODIS-derived POC are even smaller with a median difference very close to zero (−0.54 or −0.33 mg m^{-3}, respectively) and *MdAPD* <6%. For both comparisons of pairs of ocean color sensors, more significant differences in POC are typically observed in waters with POC that is relatively high compared to low POC levels in subtropical gyres, which is observed at temperate latitudes as well as in some upwelling and coastal or shelf environments. In addition, these differences appear to be enhanced in the month of July compared to January. These results suggest that a more detailed inter-sensor comparison of POC product derived from different ocean color satellite missions will require special attention paid to the analysis of differences occurring at sub-global scales of ocean regions, basins, or biogeochemical provinces.

4. Summary and future perspectives

This study describes the performance assessment and validation of a new suite of global hybrid POC algorithms for SeaWiFS, MODIS, and VIIRS-SNPP satellite ocean color sensors, which were described recently in SJR2022 (Stramski et al., 2022). The analysis of predictive capabilities of the hybrid algorithm in comparison to two other ocean color POC algorithms, i.e., the current standard global Band Ratio Power Function (BR-PF) algorithm and the Color-Index (CI) algorithm described previously in the literature, is also presented. For these analyses, we assembled two main types of validation datasets, i.e., in situ validation dataset and satellite validation datasets.

The assembled in situ validation dataset contains coincident field measurements of POC and $R_{rs}(\lambda)$ and is independent of in situ data that were used in the development of algorithms. The in situ validation dataset includes measurements made with both in-water radiometry (IWR) and above-water radiometry (AWR) approaches. Overall this dataset covers a broad range of oceanic environments including both the continental shelf and off-shelf environments. The analysis based on this in situ validation dataset demonstrated improved performance of the satellite sensor-specific hybrid POC algorithms in comparison to BR-PF and CI algorithms. For example, the median absolute percentage difference (*MdAPD*) between POC derived from the hybrid algorithms and measured POC for the entire in situ validation dataset is 18–19% and the median ratio (*MdR*) of algorithm-derived to measured POC is very close to 1 (to within 2%). The validation analysis with the independent in situ dataset serves primarily to assess the performance of the algorithms themselves irrespective of various additional sources of uncertainties that are present when the algorithms are applied to satellite observations.

The second type of validation dataset, referred to as the satellite validation dataset, serves to assess the performance of algorithms in the context of satellite applications. We assembled three satellite validation datasets, which include the satellite-in situ matchup data for the SeaWiFS, MODIS-Aqua, and VIIRS-SNPP satellite missions. The three sensor-specific satellite validation datasets were assembled using a matchup screening process based on a number of strict inclusion/exclusion criteria as detailed in Section 2.1.2. Although these validation datasets are limited in size (from <100 to over 300 data points depending on the ocean color mission), they cover a broad range of POC and provide a reasonably good basis for the assessment of uncertainties in satellite-derived POC product. The validation analysis of satellite-in situ matchups reveals deviations between the satellite-derived POC and field measurements of POC that are associated with multiple sources of uncertainty. In addition to the algorithm itself, the sources of uncertainty include satellite radiometric measurements, atmospheric correction, spatial/temporal mismatches in the validation matchup dataset, and potential errors in field measurements.

The key result from the analysis of satellite-in situ data matchups is similar to that obtained with the independent in situ validation dataset, namely the hybrid POC algorithms exhibit a clear tendency for improved performance in comparison to the BR-PF and CI algorithms. As expected, however, the performance statistics obtained from the analysis of satellite validation datasets declined to a certain degree compared to those obtained with the in situ validation dataset. For example, the *MdAPD* is in the range of about 19–27% and *MdR* in the range of 1.01–1.09 for the three sensor-specific hybrid algorithms evaluated with the satellite validation datasets. Importantly, however, we also found improvements in the performance statistics when the satellite validation datasets were additionally constrained by including only the data that are consistent with seawater bio-optical properties representative of open-ocean pelagic environments and the bottom depth criterion ($\geq$ 50 m). These criteria were used in SJR2022 for the development of global hybrid algorithms and hence, as might have been expected, their application to the satellite validation datasets resulted in improved performance statistics of the algorithms.

The investigation of potential presence of long-term trends in oceanic stock of POC at the regional, basin or global scales requires the consistency of POC data product derived from different satellite missions that provide operational continuity over multiple decades. In this study, the SeaWiFS, MODIS, and VIIRS-SNPP sensor-specific hybrid POC algorithms have shown generally good performance. When tested with independent in situ and satellite validation datasets, the median percentage difference between the algorithm-derived and measured POC was below 30%, which represents the previously recommended accuracy threshold for the standard global satellite-derived chlorophyll-a concentration product (GCOS, 2011). Additionally, the POC retrievals from the three satellite sensor-specific hybrid algorithms showed very good agreement when tested with the in situ validation dataset. High degree of consistency between the POC retrievals from the three sensor-specific algorithms was also supported when tested with satellite validation datasets, which included approximately coincident measurements made by different satellite ocean color sensors during their mission overlap.

The presented validation analyses suggest that the new suite of satellite sensor-specific hybrid POC algorithms has the potential to better represent the spatial and temporal variability within a broader range of POC than the predecessor global algorithms and to provide an improved capability for generating a consistent long-term data record of POC from multiple satellite missions. Whereas the present study is limited to the analysis of algorithms for the SeaWiFS, MODIS, and VIIRS-SNPP ocean color sensors, there is a need for further validation of algorithms specific to other ocean color sensors (see Table 4 in Stramski et al., 2022) including the use of merged satellite data products of R_{rs} (e.g., the ESA Ocean Color Climate Change Initiative, OC-CCI). One of the next research tasks is to conduct more extensive inter-sensor comparative

analysis with the use of larger satellite datasets than in the present study in order to cover the entire overlapping periods between different satellite missions. Such analysis will support further investigation of the long-term consistency of POC product derived from multiple satellite missions.

Another potential direction of research that calls for attention is related to the measurement methodology of POC, which naturally has implications to the development and interpretation of POC algorithms and derived POC product. The POC measurement protocols have been recently described in IOCCG Protocol Series (2021) and some methodological aspects in the context of development of global hybrid POC algorithms were addressed in SJR2022. Specifically, the main suite of hybrid algorithms was developed with POC data obtained with a measurement protocol that does not include the correction for DOC adsorption on sample filters used to retain the particulate matter. The validation results presented in this study are focused on these hybrid algorithms and such POC data. However, SJR2022 also provided preliminary hybrid algorithms developed with POC measurements that were subject to retroactive DOC correction based on the volume of filtered samples (Novak et al., 2018). The interpretive differences and potential advantages or disadvantages of the use of POC obtained with or without DOC correction, especially in the context of optically-based POC algorithms, were also addressed in SJR2022. Although the provided algorithm formulations involving the POC with DOC correction are preliminary, a limited validation analysis of these algorithms was conducted in this study using the independent in situ validation dataset and satellite validation datasets. The results of this analysis are presented in Supplementary Material, which can provide a useful benchmark for further work as more DOC-corrected measurements of POC are collected along with optical measurements across diverse oceanic environments.

Credit author statement

Conceptualization: IDJ, DS, RAR.
Methodology: IDJ, DS, RAR.
Data acquisition and curation: IDJ, DHR.
Formal data analysis: IDJ, DS, RAR.
Writing original draft: IDJ, DS.
Revising the manuscript: IDJ, DS, RAR, DHR.
Approval of the manuscript: IDJ, DS, RAR, DHR.

Declaration of Competing Interest

Ishan D. Joshi reports financial support was provided by NASA. Dariusz Stramski reports financial support was provided by NASA. Rick A. Reynolds reports financial support was provided by NASA. Dale H. Robinson reports financial support was provided by NASA.

Data availability

Data will be made available on request.

Acknowledgments

The study was supported by the National Aeronautics and Space Administration (NASA Grants 80NSSC18K0956 and 80NSSC21K0700). We thank the NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS: https://seabass.gsfc.nasa.gov/), NSF Biological and Chemical Oceanography Data Management Office (BCO-DMO: https://www.bco-dmo.org/), PANGAEA (https://www.pangaea.de/), the NOAA National Center for Environmental Information (NCEI: https://www.nodc.noaa.gov) and the NOAA CoastWatch West Coast Node (https://coastwatch.pfeg.noaa.gov) for providing access to data used in this study, as well as all scientists who collected measurements and contributed to these archives. We acknowledge the NOAA Southwest Fisheries Science Center (https://www.fisheries.noaa.gov/about/southwest-fisheries-science-center) for providing computing resources. We also thank the NASA Ocean Biology Processing Group for the production and distribution of the SeaWIFS, MODIS-Aqua, and VIIRS-SNPP imagery used to assemble the satellite validation datasets. We thank the Associate Editor, Frédéric Mélin, and three anonymous reviewers for valuable comments on the manuscript.

Appendix A. Supplementary Material

Supplementary Material to this article can be found online at https://doi.org/10.1016/j.rse.2022.113417.

References

Allison, D.B., Stramski, D., Mitchell, B.G., 2010. Empirical ocean color algorithms for estimating particulate organic carbon in the Southern Ocean. J. Geophys. Res. 115, C10044. https://doi.org/10.1029/2009JC006040.

Amante, C., Eakins, B.W., 2009. ETOPO1 1 Arc-minute global relief model: Procedures, data sources, and analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5C8276M.

Austin, R.W., 1974. The remote sensing of spectral radiance from below the ocean surface. In: Jerlov, N.G., Nielsen, E.S. (Eds.), Optical Aspects of Oceanography. Academic Press, London, pp. 317–344.

Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ. 102, 12–23. https://doi.org/10.1016/j.rse.2006.01.015.

Behrenfeld, M.J., Boss, E., Siegel, D.A., Shea, D.M., 2005. Carbon-based ocean productivity and phytoplankton physiology from space. Glob. Biogeochem. Cycles 19, GB1006. https://doi.org/10.1029/2004GB002299.

Bianchi, T.S., Bauer, J.E., 2011. Particulate organic carbon cycling and transformation. In: Wolanski, E., McLusky, D.S. (Eds.), , Treatise on Estuarine and Coastal Science, Vol 5. Academic Press, Waltham, pp. 69–117.

Boyd, P.W., Newton, P.P., 1999. Does planktonic community structure determine downward particulate organic carbon flux in different oceanic provinces? Deep-Sea Res. Part I: Oceanogr. Res. Papers 46, 63–91. https://doi.org/10.1016/S0967-0637(98)00066-1.

Brewin, R.J.W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H., Lee, Z., Maritorena, S., Mélin, F., Peters, M., Platt, T., Regner, P., Smyth, T., Steinmetz, F., Swinton, J., Werdell, J., White, G.N., 2015. The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms. Remote Sens. Environ. 162, 271–294. https://doi.org/10.1016/j.rse.2013.09.016.

Broomell, S.B., Budescu, D.V., Por, H.-H., 2011. Pair-wise comparisons of multiple models. Judgm. Decis. Mak. 6 (8), 821–831.

Cetinić, I., Perry, M.J., Briggs, N.T., Kallin, E., D'Asaro, E.A., Lee, C.M., 2012. Particulate organic carbon and inherent optical properties during 2008 North Atlantic bloom experiment. J. Geophys. Res. 117, C06028. https://doi.org/10.1029/2011JC007771.

Collos, Y., Jauzein, C., Hatey, E., 2014. Particulate carbon and nitrogen determinations in tracer studies: the neglected variables. Appl. Radiat. Isot. 94, 14–22. https://doi.org/10.1016/j.apradiso.2014.06.015.

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., Roy, S., Stramski, D., Thomalla, S., Platt, T., Sathyendranath, S., 2017. Validation and intercomparison of ocean color algorithms for estimating particulate organic carbon in the oceans. Front. Mar. Sci. 4, 251. https://doi.org/10.3389/fmars.2017.00251.

Fenchel, T., 1988. Marine plankton food chains. Ann. Rev. Ecol. Syst. 19, 19–38. https://doi.org/10.1146/annurev.es.19.110188.000315.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281, 237–240. https://doi.org/10.1126/science.281.5374.237.

Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O., Huang, X., Davis, A.B., Kalashnikova, O., Thompson, D.R., Remer, L.A., Boss, E., Coddington, O., Deschamps, P.-Y., Gao, B.-C., Gross, L., Hasekamp, O., Omar, A., Pelletier, B., Ramon, D., Steinmetz, F., Zhai, P.-W., 2019. Atmospheric correction of satellite ocean-color imagery during the PACE era. Front. Earth Sci. 7, 145. https://doi.org/10.3389/feart.2019.00145.

Gardner, W.D., Richardson, M.J., Carlson, C.A., Hansell, D., Mishonov, A.V., 2003. Determining true particulate organic carbon: bottles, pumps and methodologies. Deep-Sea Res. II Top. Stud. Oceanogr. 50, 655–674. https://doi.org/10.1016/S0967-0645(02)00589-1.

Gardner, W.D., Mishonov, A.V., Richardson, M.J., 2006. Global POC concentrations from in-situ and satellite data. Deep-Sea Res. II Top. Stud. Oceanogr. 53, 718–740. https://doi.org/10.1016/j.dsr2.2006.01.029.

GCOS, 2011. Systematic observation requirements from satellite-based data products for climate 2011 Update. Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update)". Technical Report, No. 154. World Meteorological Organisation (WMO), Geneva 2, Switzerland.

Gordon, H.R., McCluney, W.R., 1975. Estimation of the depth of sunlight penetration in the sea for remote sensing. Appl. Opt. 14, 413–416. https://doi.org/10.1364/AO.14.000413.

Hu, C., Carder, K.L., Muller-Karger, F.E., 2001. How precise are SeaWiFS Ocean color estimates? Implications of digitization-noise errors. Remote Sens. Environ. 76, 239–249. https://doi.org/10.1016/S0034-4257(00)00206-6.

Hu, C., Lee, Z., Franz, B., 2012. Chlorophyll a algorithms for oligotrophic oceans: a novel approach based on three-band reflectance difference. J. Geophys. Res. 117, C01011. https://doi.org/10.1029/2011JC007395.

Hu, C., Feng, L., Lee, Z., 2013. Uncertainties of SeaWiFS and MODIS remote sensing reflectance: implications from clear water measurements. Remote Sens. Environ. 133, 168–182. https://doi.org/10.1016/j.rse.2013.02.012.

IOCCG Protocol Series, 2019. Protocols for satellite ocean colour data validation: in situ optical radiometry. In: Zibordi, G., Voss, K.J., Johnson, B.C., Mueller, J.L. (Eds.), IOCCG Ocean Optics and Biogeochemistry Protocols for Satellite Ocean ColourSensor Validation, Volume 3.0. IOCCG, Dartmouth, NS, Canada. https://doi.org/10.25607/OBP-691.

IOCCG Protocol Series, 2021. Particulate organic matter sampling and measurement protocols: consensus towards future ocean color missions. In: Chaves, J.E., Cetinić, I., Estapa, M., Gardner, W., Gōni, M., Graff, J.R., Hernes, P., Lam, P.J., Lomas, M.W., Mannino, M., Novak, M.G., Turnewitsch, R., Werdell, P.J., Westberry, T.K., Liu, Z. (Eds.), IOCCG Ocean Optics and Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation, Volume 6.0. IOCCG, Dartmouth, NS, Canada. https://doi.org/10.25607/OBP-1646.

Joshi, I.D., D'Sa, E.J., 2020. Optical properties using adaptive selection of NIR/SWIR reflectance correction and quasi-analytic algorithms for the MODIS-Aqua in estuarine-ocean continuum: application to the northern Gulf of Mexico. IEEE Trans. Geosci. Remote Sens. 58, 6088–6105. https://doi.org/10.1109/TGRS.2020.2973157.

Knap, A.H., Michaels, A., Close, A.R., Ducklow, H.W., Dickson, A.G., 1996. Protocols for the Joint Global Ocean Flux Study (JGOFS) core measurements. JGOFS, Reprint of the IOC Manuals and Guides No. 29, UNESCO 1994, 19, 210 pp.

Le, C., Zhou, X., Hu, C., Lee, Z., Li, L., Stramski, D., 2018. A color-index-based empirical algorithm for determining particulate organic carbon concentration in the ocean from satellite observations. J. Geophys. Res. Oceans 123, 7407–7419. https://doi.org/10.1029/2018JC014014.

Liu, H., Li, Q., Bai, Y., Yang, C., Wang, J., Zhou, Q., Hu, S., Shi, T., Liao, X., Wu, G., 2021. Improving satellite retrieval of oceanic particulate organic carbon concentrations using machine learning methods. Remote Sens. Environ. 256, 112316 https://doi.org/10.1016/j.rse.2021.112316.

Liu, H., Shah, S., Jiang, W., 2004. On-line outlier detection and data cleaning. Comp. Chem. Eng. 28, 1635–1647. https://doi.org/10.1016/j.compchemeng.2004.01.009.

Loisel, H., Nicolas, J.-M., Deschamps, P.-Y., Frouin, R., 2002. Seasonal and inter-annual variability of particulate organic matter in the global ocean. Geophys. Res. Lett. 29 (24), 2196. https://doi.org/10.1029/2002GL015948.

Longhurst, A.R., Harrison, W.G., 1989. The biological pump: profiles of plankton production and consumption in the upper ocean. Prog. Oceanogr. 22, 47–123. https://doi.org/10.1016/0079-6611(89)90010-4.

Mishonov, A.V., Gardner, W.D., Richardson, M.J., 2003. Remote sensing and surface POC concentration in the South Atlantic. Deep-Sea Res. II Top. Stud. Oceanogr. 50, 2997–3015. https://doi.org/10.1016/j.dsr2.2003.07.007.

Mobley, C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. Appl. Opt. 38, 7442–7455. https://doi.org/10.1364/AO.38.007442.

Moran, S.B., Charette, M.A., Pike, S.M., Wicklund, C.A., 1999. Differences in seawater particulate organic carbon concentration in samples collected using small- and large-volume methods: the importance of DOC adsorption to the filter blank. Mar. Chem. 67, 33–42. https://doi.org/10.1016/S0304-4203(99)00047-X.

Morel, A., Huot, Y., Gentili, B., Werdell, P.J., Hooker, S.B., Franz, B.A., 2007. Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. Remote Sens. Environ. 111, 69–88. https://doi.org/10.1016/j.rse.2007.03.012.

Morley, S., Henderson, M.G., Woodroffe, J.R., Brito, T.V., 2016. The global positioning system constellation as a space weather monitor: comparison of electron measurements with Van Allen probes data. Space Weather 14, 76–92. https://doi.org/10.1002/2015SW001339.

Morley, S.K., Brito, T.V., Welling, D.T., 2018. Measures of model performance based on the log accuracy ratio. Space Weather 16, 69–88. https://doi.org/10.1002/2017SW001669.

Novak, M.G., Cetinić, I., Chaves, J.E., Mannino, A., 2018. The adsorption of dissolved organic carbon onto glass fiber filters and its effect on the measurement of particulate organic carbon: a laboratory and modeling exercise. Limnol. Oceanogr. Methods 16, 356–366. https://doi.org/10.1002/lom3.10248.

Pabi, S., Arrigo, K.R., 2006. Satellite estimation of marine particulate organic carbon in waters dominated by different phytoplankton taxa. J. Geophys. Res. 111, C09003. https://doi.org/10.1029/2005JC003137.

Pahlevan, N., Sarkar, S., Franz, B.A., 2016. Uncertainties in coastal ocean color products: impacts of spatial sampling. Remote Sens. Environ. 181, 14–26. https://doi.org/10.1016/j.rse.2016.03.022.

Pearson, R.K., Neuvo, Y., Astola, J., Gabbouj, M., 2016. Generalized Hampel filters. EURASIP J. Adv. Signal Process. 87 (2016) https://doi.org/10.1186/s13634-016-0383-6.

Rasse, R., Dall'Olmo, G., Graff, J., Westberry, T.K., van Dongen-Vogels, V., Behrenfeld, M.J., 2017. Evaluating optical proxies of particulate organic carbon across the surface Atlantic Ocean. Front. Mar. Sci. 4, 367. https://doi.org/10.3389/fmars.2017.00367.

Reynolds, R.A., Stramski, D., Neukermans, G., 2016. Optical backscattering of particles in Arctic seawater and relationships to particle mass concentration, size distribution, and bulk composition. Limnol. Oceanogr. 61, 1869–1890. https://doi.org/10.1002/lno.10341.

Seegers, B.N., Stumpf, R.P., Schaeffer, B.A., Loftin, K.A., Werdell, P.J., 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. Opt. Express 26, 7404–7422. https://doi.org/10.1364/OE.26.007404.

Son, Y.B., Gardner, W.D., Mishonov, A.V., Richardson, M.J., 2009. Multispectral remote-sensing algorithms for particulate organic carbon (POC): the Gulf of Mexico. Remote Sens. Environ. 113, 50–61. https://doi.org/10.1016/j.rse.2008.08.011.

Stramska, M., Stramski, D., 2005. Variability of particulate organic carbon concentration in the north polar Atlantic based on ocean color observations with Sea-viewing Wide Field-of-view Sensor (SeaWiFS). J. Geophys. Res. 110, C10018. https://doi.org/10.1029/2004JC002762.

Stramska, M., Cieszyńska, A., 2015. Ocean colour estimates of particulate organic carbon reservoirs in the global ocean – revisited. Int. J. Remote Sens. 35, 3675–3700. https://doi.org/10.1080/01431161.2015.1049380.

Stramski, D., Reynolds, R.A., Kahru, M., Mitchell, B.G., 1999. Estimation of particulate organic carbon in the ocean from satellite remote sensing. Science 285, 239–242. https://doi.org/10.1126/science.285.5425.239.

Stramski, D., Reynolds, R.A., Babin, M., Kaczmarek, S., Lewis, M.R., Röttgers, R., Sciandra, A., Stramska, M., Twardowski, M.S., Franz, B.A., Claustre, H., 2008. Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic oceans. Biogeosciences 5, 171–201. https://doi.org/10.5194/bg-5-171-2008.

Stramski, D., Joshi, I., Reynolds, R.A., 2022. 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. 269, 112776 https://doi.org/10.1016/j.rse.2021.112776.

Świrgoń, M., Stramska, M., 2015. Comparison of in situ and satellite ocean color determinations of particulate organic carbon concentration in the global ocean. Oceanologia 57, 25–31. https://doi.org/10.1016/j.oceano.2014.09.002.

Tran, T.K., Duforêt-Gaurier, L., Vantrepotte, V., Jorge, D.S.F., Mériaux, X., Cauvin, A., Fanton d'Andon, O., Loisel, H., 2019. Deriving particulate organic carbon in coastal waters from remote sensing: inter-comparison exercise and development of a maximum band-ratio approach. Remote Sens. 11 (23), 2849. https://doi.org/10.3390/rs11232849.

Turnewitsch, R., Springer, B.M., Kiriakoulakis, K., Vilas, J.C., Arístegui, J., Wolff, G., Peine, F., Werk, S., Graf, G., Waniek, J.J., 2007. Determination of particulate organic carbon (POC) in seawater: the relative methodological importance of artificial gains and losses in two glass-fiber-filter-based techniques. Mar. Chem. 105, 208–228. https://doi.org/10.1016/j.marchem.2007.01.017.

Volk, T., Hoffert, M.I., 1985. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes. In: Sundquist, E.T., Broecker, W.S. (Eds.), The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present, Geophys. Monogr. Ser, vol. 32, AGU, Washington, D. C, pp. 99–110.

Wang, F., Wang, Y., Chen, Y., Liu, K., 2020. Remote sensing approach for the estimation of particulate organic carbon in coastal waters based on suspended particulate concentration and particle median size. Mar. Pollut. Bull. 158, 111382 https://doi.org/10.1016/j.marpolbul.2020.111382.