



Evaluation of Chlorophyll-a estimation using Sentinel 3 based on various algorithms in southern coastal Vietnam



Nguyen An Binh ^{a,b,*}, Pham Viet Hoa ^a, Giang Thi Phuong Thao ^a, Ho Dinh Duan ^a, Phan Minh Thu ^c

^a Ho Chi Minh City Institute of Resources Geography, Vietnam Academy of Science and Technology, Ho Chi Minh City, Viet Nam

^b Graduate University of Science and Technology, Ha Noi, Viet Nam

^c Institute of Oceanography, Vietnam Academy of Science and Technology, Khanh Hoa, Viet Nam

ARTICLE INFO

Keywords:

OLCI
Sentinel 3
Bio-optical algorithms
DINEOF
C2RCC
DSF
Chlorophyll-a

ABSTRACT

This paper aims to assess the potential of Ocean Land Colour Instrument (OLCI) for the retrieval of chlorophyll-a (chl-a) over southern coastal waters of Vietnam. For that purpose, four chlorophyll-a ocean color (OC) algorithms (OC4ME and three new OC version 7 OC4, OC5, OC6) were applied based on water-leaving reflectance obtained from two atmospheric correction processors (C2RCC and DSF). To overcome high cloud coverage in the area of interest, full spatial data reconstruction was implemented using Data Interpolating Empirical Orthogonal Functions (DINEOF). Numerical error metrics of in situ measurements ($n = 49$) collected in different ship-based campaigns has been assessed for Sentinel-3A (S-3A) and 3B (S-3B) as well as on the combined products built from these two later satellites. Results showed that products based on C2RCC significantly outperformed DSF. For chl-a algorithms, C2RCC-based OC5 gave the most accurate retrieval while applied to S-3A ($R^2: 0.58$, RMSE: 1.018 mg m^{-3} , MAPE: 49.4 %), S-3B ($R^2: 0.75$, RMSE: 0.776 mg m^{-3} , MAPE: 37.3 %), and synergy datasets ($R^2: 0.70$, RMSE: 0.844 mg m^{-3} , MAPE: 42.5 %). With >50 % of observations missing due to cloud cover, DINEOF provides a promising solution to reconstruct the full spatial information. The successfully demonstrated retrieval of chl-a in our study presents potential for daily monitoring when combining observations from S-3A/B to further improve our understanding of the spatio-temporal dynamics of coastal ecosystems.

1. Introduction

Remotely sensed records allow us to have comprehensive insights into aquatic-related processes, such as primary production and harmful algal bloom (IOCCG, 2021), eutrophication (Kitsiou and Karydis, 2011), as well as analysis of the variation and trends of different biological variables (Alvera-Azcarate et al., 2016; Ji et al., 2018; Loisel et al., 2017, 2014). Chlorophyll-a (chl-a), one of the most crucial ocean variables, is mainly used for phytoplankton pigment assessment and an indirect trophic signal (Ha et al., 2013; Mercado et al., 2016; O'Reilly and Werdell, 2019). Due to the oxygen manufacturing processes based on photosynthesis in aquatic organisms (absorb visible sunlight), optical satellites can take the advantage to identify through ocean color, enabling the capacity for mapping the entire ocean's primary productivity (Falkowski, 2012).

The first spaceborne-based chl-a algorithm combined two equations from two bands ratio in order to retrieve the concentration of

phytoplankton pigments was applied in the pioneer Coastal zone color scanner (CZCS) sensor (Gordon et al., 1983). However, inconsistency while merging data from two equations was emphasized in (O'Reilly et al., 1998). To resolve this issue, a new ocean color chl-a (OC) algorithm based on a single fourth-order polynomial equation has been proposed for the first application in Sea-viewing Wide Field-of-view Sensor (SeaWiFS) (O'Reilly et al., 1998). Different version OC algorithms were updated such as OC2, OC3, OC4 (O'Reilly and Maritorena, 2000), OC4ME (Morel, 2007), OC5 (Gohin et al., 2002), and by NASA bio-Optical Marine Algorithm Dataset (NOMAD) (Werdell and Bailey, 2005) with SeaBASS datasets in 2009 for numerical optical satellites, available in https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/. The newest version 7 with the new term of OC6 was tested globally in 25 sensors (O'Reilly and Werdell, 2019).

The recent Ocean Land Colour Instrument (OLCI) onboard Sentinel-3 enables global daily ocean color images combining both satellite 3A (February – 2016) and 3B (April – 2018) at full resolution of 300 m

* Corresponding author at: Ho Chi Minh City Institute of Resources Geography, Vietnam Academy of Science and Technology, Ho Chi Minh City, Viet Nam.
E-mail address: nabinh@hcmig.vast.vn (N.A. Binh).

(Donlon et al., 2012). Recent studies have assessed the performance of OLCI-based OC algorithms at a global scale (O'Reilly and Werdell, 2019), for case-1 open ocean water (Tilstone et al., 2021), and coastal areas (Kratzer and Plowey, 2021; Lavigne et al., 2021; Pahlevan et al., 2020). Nevertheless, the applications of OLCI are still needed for more validation (Lavigne et al., 2021).

One of the biggest challenges for mapping ocean color from space observations is atmospheric corrections. The common assumption over open ocean waters is the negligible water reflectance signal in the Near Infrared, NIR (Gordon and Wang, 1994). However, non-zero NIR water reflectance often appears in coastal areas, and numerical approaches have been suggested to deal with the issues. For instance, machine learning approaches have been developed for different sensors such as the Case 2 Regional CoastColour (C2RCC) model based on artificial neural network (Brockmann et al., 2016; Doerffer and Schiller, 2007). Taking advantage of wavelength in optical instruments, ShortWave InfraRed (SWIR) band (1020 nm) in the next generation Sentinel 3/OLCI (2016 for S-3A, 2018 for S-3B) can be used for accurate atmospheric correction (Vanhellemont and Ruddick, 2021). A new adaptive version of Acolite using SWIR bands in OLCI for Dark spectrum fitting (DSF) processors showed the most accurate atmospheric correction in 490–681 nm and performed less noisy than L2-WFR, POLYMER, C2RCC, SeaDAS in Belgian coastal waters (Vanhellemont and Ruddick, 2021).

Atmospheric correction schemes and bio-optical algorithms are essential for reliable retrieval of chl-a using remotely sensed data. Meanwhile, gap-filling tools enable the potency of full spatial reconstruction of datasets overcoming cloud cover and noisy pixels. In optical remote sensing of ocean color, Data Interpolating Empirical Orthogonal Functions (DINEOF) has been successfully applied to different ocean variables, such as sea surface temperature (SST) (Alvera-Azcárate et al., 2005), chl-a (Ji et al., 2018; Li and He, 2014), total suspended matter (TSM) (Nechad et al., 2011), turbidity (Alvera-Azcárate et al., 2015), sea surface salinity (SSS) (Alvera-Azcárate et al., 2016), and harmful algal (Anderson et al., 2016). Compared to the most widely used DINEOF, other methods such as ensemble optimal interpolation (Oke et al., 2010), numerical models (Konik et al., 2019), and the deep learning-based Data-Interpolating Convolutional Auto-Encoder (DINCAE) with updated version 2.0 (Barth et al., 2022) and recent applications for chl-a reconstruction (Han et al., 2020; Ji et al., 2021).

With more than 3000 km of coastline, the Vietnamese coastal region plays a vital role in the nation's development and socio-economic goals. Yet, research on ocean color remote sensing has recently started with few studies published in the past decade. (Ha et al., 2013) estimated chl-a based on the green and blue band reflectance ratio over tropical coastal waters in northern Vietnam using Moderate Resolution Imaging Spectroradiometer (MODIS) data. (Loisel et al., 2017) assessed the performance of six different bio-optical chl-a algorithms and analyzed a time-series of 10 years (2002–2012) chl-a variability over the whole Vietnamese coastal waters through Medium Resolution Imaging Spectrometer (MERIS). Under tropical monsoon climate, high cloud cover is prevalent across coastal Vietnam, resulting in the breakdown of mapping entire ocean variables and the lack of match-up exercises for validation against field measurement (Loisel et al., 2017; Ngoc et al., 2019, 2020).

This study aims to evaluate the performance of Sentinel 3 chl-a algorithms over Southern coastal Vietnam. Our approaches used multi-processing steps to retrieve chl-a at the full spatial distribution. Atmospheric corrections are performed using the C2RCC and DSF processors. For chl-a algorithms, the standard algorithm OC4ME and new OC version 7 (OC4, OC5, OC6) were applied to water-leaving reflectance products. For the post-processing, DINEOF was used for the reconstruction of missing data. Cross-validation of DINEOF and in-situ field measurements from different campaigns were used to assess the performance of the models. The study opens up novel discussions on the capacity of OLCI observations and DINEOF to retrieve gap-free chl-a spatio-temporal patterns across the Vietnamese coastal area while

overcoming the regions with high-frequency cloud cover.

2. Description of the study areas and the different in situ and satellite data

2.1. Study area

The study area is located in southern coastal Vietnam ($11^{\circ}20' - 12^{\circ}53'$), covering an approximately 5848.7 km^2 along the coastal line of nearly 600 km (Fig. 1). Social-economic development over the site focus on three bays: Van Phong (VP) and Nha Trang (NT), located in Khanh Hoa province, and Phan Rang (PR) belongs to Ninh Thuan province.

2.2. Field campaigns

Data from five field surveys were collected in different periods and locations. There were four campaigns organized in VP and NT in 2018 and 2019. An additional field survey occurred in PR in 2018. The description of total 49 field samples is provided in Table 1.

Water samples were collected under 1 m from the surface mixed layer using a Niskin bottle. The filtered water samples were then stored in dark cooling box (0°C) before returning to the laboratory in 2–6 h. For the laboratory analysis, chl-a concentration was calculated by corrected chlorophyll-a method (Jeffrey et al., 1997).

2.3. Remote sensing data

Full resolution (300 m) of Sentinel 3A (S-3A) and 3B (S-3B) were downloaded from the ESA hub. The onboard sensor OLCI contains 21 bands, with wavelengths ranging from 400 to 1020 nm. There are two periods of image collection periods based on different field campaigns: 46 images (S-3A: 22, S-3B: 24) from 1/10/2018 to 30/11/2018 and 56 images (S-3A: 26, S-3B: 30) from 1/7/2019 to 31/8/2019.

3. Methods

3.1. Atmospheric correction

C2RCC and DSF were applied in the atmospheric correction schemes. For C2RCC, Sentinel Application Platform (SNAP) was used for further processed with default settings. The masking phase includes: quality_flags_land, quality_flags_bright, quality_flags_invalid, Rotsa_OOR, Rho_w_OOR, Cloud_risk, Iop_OOR. For DSF, a new beta version of Acolite with updated for OLCI processing was download in <https://github.com/acolite/acolite>. The band 1020 nm was used for non-water masking with threshold of 0.05. Wavelength to check for negative reflectance ranging from 400 to 900 nm.

3.2. Chlorophyll-a retrieve algorithms

Chl-a empirical algorithms based on fourth-order polynomial were assessed in the study, following the Eq. (1) and (2) (O'Reilly et al., 1998):

$$\log_{10}\text{Chla} = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4 \quad (1)$$

where:

$$X = \log_{10} \left[\max \left(\frac{R_{rs}(\text{blue})}{R_{rs}(\text{green})} \right) \right] \quad (2)$$

with $R_{rs}(\text{blue})$ and $R_{rs}(\text{green})$ represent the 413, 443, 490, 510 nm and 560 nm for OLCI, respectively. The coefficients $a_0 - a_4$ were empirically fitted based on specific sensors and different bio-optical algorithms.

Four OC chl-a algorithms were applied in this study including OC4ME, OC4, OC5, and OC6. Among them, OC4ME is the standard algorithms that followed ESA documents, and three recent OC version 7

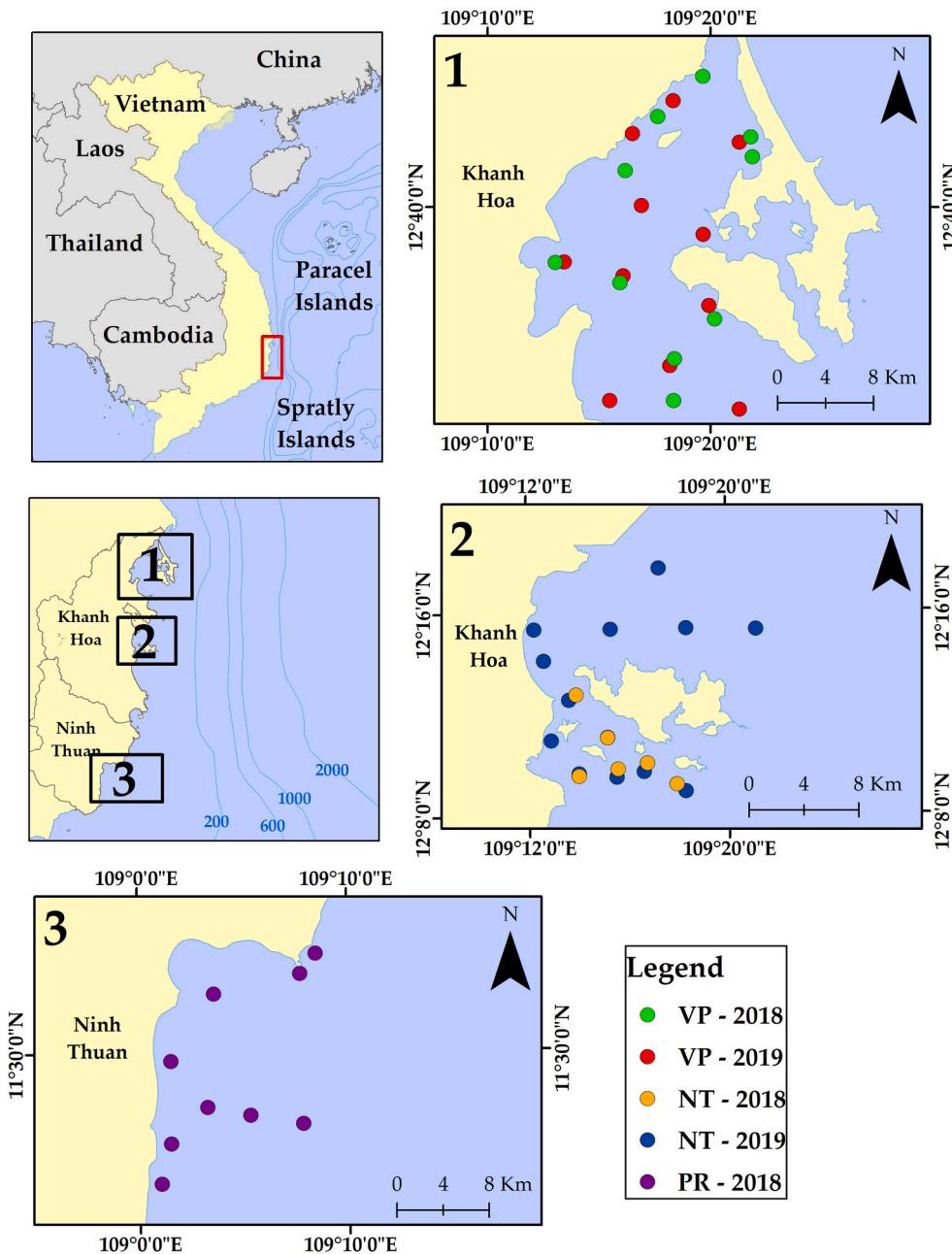


Fig. 1. Location of the study area and field sampling in (1) Van Phong Bay, (2) Nha Trang Bay, and (3) Phan Rang Bay.

Table 1
Statistical description of chl-a (mg m^{-3}) data collection.

Location	Period	N	min	max	median
VP	July 31th, 2019	11	0.568	5.450	1.059
NT	August 2nd, 2019	13	0.252	2.331	0.546
VP	October 31th, 2018	10	0.675	8.372	2.077
NT	November 5th, 2018	6	0.478	1.353	0.838
PR	October 25th, 2018	9	0.576	1.602	1.209

calibrated for OLCI were also used (O'Reilly and Werdell, 2019). The terms 4,5,6 represent the number of bands used in the algorithms. Table 2 provides the key wavelengths and coefficients in each algorithm. Note that OC6 used the mean reflectance values of 560 nm and 665 nm in the denominator, as taking the band at approximately 613 nm in the model.

3.3. Data reconstruction

To establish the DINEOF process, the original datasets are stored in a spatio-temporal matrix with $y \times x \times t$ dimension (x and y are the spatial dimensions of latitude and longitude respectively, and t is the number of images in each dataset representing the temporal dimension). Based on Empirical Orthogonal Function (EOF) decomposition, the first EOF mode is calculated by using Singular Value Decomposition (SVD) technique. Once the convergence is reached, the procedure is repeated with the next 2,3,4..., n EOF mode. By applying cross-validation with 3 % of reconstructed pixels compared against initial pixels, the total number of EOF modes is calculated. Finally, the optimal EOF mode (lowest root-mean-square error RMSE) will be used for whole pixel reconstruction over each image in the dataset. A full description of methods can be found in (Beckers and Rixen, 2003) and (Alvera-Azcárate et al., 2005).

DINEOF was applied to 48 different chl-a datasets separated by two

Table 2

Coefficients for Bio-optical algorithms for OLCI/Sentinel-3.

Algorithms	Blue	Green	a0	a1	a2	a3	a4
OC4ME	443 > 490 > 510	560	0.45027	-3.259491	1.9743	3.522731	0.949586
OC4	443 > 490 > 510	560	0.42540	-3.21679	2.86907	-0.62628	-1.09333
OC5	413 > 443 > 490 > 510	560	0.43213	-3.13001	3.05479	-1.45176	-0.24947
OC6	413 > 443 > 490 > 510	mean (560,665)	0.95039	-3.05404	2.17992	-1.12097	0.15262

periods (2018, 2019), four OC algorithms (OC4ME, OC4, OC5, and OC6), two atmospheric corrections (C2RCC and DSF), and two satellites (S-3A, S-3B, and S-3A + S-3B).

3.4. Statistical analysis

3.4.1. Cross-validation of DINEOF

We validated the DINEOF reconstruction using cross-validation, with 3 % (automatic random selection) of original pixels were compared to the reconstructed pixels. Missing data (or percentage of cloud cover) were calculated by the percentage of missing water pixels against all water pixels over the research area. Root mean square error (RMSE) for each process was determined following the equation:

$$\text{RMSE}_{\text{DINEOF}} = \sqrt{\frac{\sum_{i=1}^n (x_i^{\text{ref}} - x_i^{\text{org}})^2}{n}} \quad (3)$$

where n is the number of cross-validation pixels, x_i^{ref} and x_i^{org} represent the chl-a values from reconstructed datasets and original datasets, respectively.

3.4.2. Model performance with in-situ measurement data

We evaluated the models using field data collected in different periods (2018, 2019) and locations (VP, NT, PR). For that purpose, daily match-up exercises were performed using 3x3 pixels following the methodology of Jamet et al. (2011). Pearson R, slope, and intercept were used to evaluate linear regression from different algorithms. Two primary metrics are root mean square errors (RMSE) and median absolute percentage error (MAPE) to measure the model accuracy. In addition, Mean absolute error (MAE) and bias were also calculated in log transformation as suggested in (Seegers et al., 2018) for better assessment of the water models.

$$R_{\text{Pearson}} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad (5)$$

$$\text{MAPE} = 100 \times \text{median}\left(\frac{|y_i - x_i|}{x_i}\right) \quad (6)$$

$$\text{MAE} = 10 \left(\frac{\sum_{i=1}^n |\log_{10} y_i - \log_{10} x_i|}{n} \right) \quad (7)$$

$$\text{bias} = 10 \left(\frac{\sum_{i=1}^n (\log_{10} y_i - \log_{10} x_i)}{n} \right) \quad (8)$$

where, n is the number of samples (49), y_i is the chl-a values estimated from satellite data and, x_i is the chl-a values measured in-situ.

We also visualize the Taylor diagram for evaluating water models based on multi aspects (i.e., the centred root mean square difference, the correlation, and the standard deviation), as detail described in (Taylor, 2001).

4. Results and analysis

4.1. Reconstruction process

Tables 3 and 4 summarize the performance of the DINEOF reconstruction using C2RCC and DSF-based products, respectively. Overall, the quality of images datasets in 2019 seems better than in 2018, with a lower percentage of missing data. Gap-filling processes in all 2019 datasets were also better than 2018 datasets, i.e., RMSE_{DINEOF} on average improvement approximately 1.5 times in C2RCC-based datasets. Cross-validation errors for all OC algorithms obtained the lowest in OC5 (C2RCC: 0.498 mg m⁻³, DSF: 0.387 mg m⁻³) and highest in OC6 (C2RCC: 0.864 mg m⁻³, DSF: 0.832 mg m⁻³). OC5 is also the best model with major better statistical metrics than other OC algorithms.

Due to the concurrent overlapping of acquisition dates in S-3A and S-3B in 2018, the combination of 2019 datasets using these two satellites resulted in a slight increase in data number. Particularly in C2RCC-based datasets, total images acquired in 2019 exhibited 49 daily records for the composite dataset, with 26 images of S-3A and 30 images of S-3B, and 7 scenes were acquired by both satellites in the same day. The higher number of images in 2019 might be explained by the interlaced sensing dates between S-3A and S-3B over the study area, resulting in images captured on different days, which helps to increase the temporal resolution for composite datasets. The percentage of missing data in DSF is just above 5 % than C2RCC in both years of 2018 and 2019 due to the masking phase of different processors (Fig. 2). For more results, see S1 in supplementary material.

4.2. Validations with in-situ data

The assessment of the chl-a estimation from the reflectance data based on C2RCC processors is summarised in Table 5. For S-3A, OC4ME and OC5 give the best performance, with an R^2 values of 0.58, and a RMSE value around 1.02 mg m⁻³. MAPE achieved by OC4ME (41.2 %) is somewhat lower than OC5 (49.4 %). There are no significant differences between the log transformation of MAE and biases in all OC, which on average, these values are 1.74 and 0.873, respectively. All algorithms have an average slope of 0.51 – 0.62, with an offset of 0.4 – 0.5, indicating underestimating, especially in high chl-a concentration areas. The biases also imply chl-a estimations are, in general, <10 – 20 % of those measured. For S-3B, OC4ME, OC4, and OC5 continue to work well as all listed error metrics. The average statistics for these algorithms include R^2 : 0.74, RMSE: 0.78 mg m⁻³, MAPE: 38.6 %. MAE and bias of four OC algorithms seem to be the same, with around 1.6 and 1.1, respectively. Notably, the difference in all S-3B water models is found only in OC6, with error metrics significantly higher than remains algorithms, i.e., RMSE higher nearly 10 – 40 % than others. By comparing between two satellites, statistics strongly suggest that S-3B-based algorithms performed better than S-3A. Among 4 OC algorithms, the best performing OC5 exhibited in both satellite 3A and 3B with the lowest RMSE 1.018 mg m⁻³ and 0.776 mg m⁻³, respectively (Figs. 3 and 4).

The observations based on DSF-derived chl-a products are provided in Table 6. For S-3A, OC6 and OC5 outperformed OC4ME and OC4, with RMSE obtained by OC5 (1.451 mg m⁻³) slightly lower than OC6 (1.586 mg m⁻³). MAPE obtained by OC5 is however better than OC6, with a significant improvement from 112.8 % to 98.9 %. Meanwhile, all OC algorithms achieved the same results in S-3B. OC6 is only the model

Table 3

Description of datasets, dimensions, missing data, and RMSE_{DINEOF} for DINEOF reconstruction from different chl-a algorithms. The atmospheric correction C2RCC.

Periods	Satellites	Dimensions	Missing data (%)	RMSE _{DINEOF} (mg m^{-3})			
				OC4ME	OC4	OC5	OC6
Oct-Nov (2018)	S-3A	179 × 537 × 21	61.17	0.815	0.786	0.785	0.864
	S-3B	179 × 537 × 24	66.25	0.795	0.805	0.769	0.797
	Total	179 × 537 × 28	60.58	0.715	0.708	0.714	0.755
July-Aug (2019)	S-3A	179 × 537 × 26	50.87	0.522	0.540	0.498	0.550
	S-3B	179 × 537 × 30	49.17	0.667	0.669	0.686	0.644
	Total	179 × 537 × 49	46.80	0.507	0.550	0.501	0.639

Table 4

Description of datasets, dimensions, missing data, and RMSE_{DINEOF} for DINEOF reconstruction from different chl-a algorithms. The atmospheric correction DSF.

Periods	Satellites	Dimensions	Missing data (%)	RMSE _{DINEOF} (mg m^{-3})			
				OC4ME	OC4	OC5	OC6
Oct-Nov (2018)	S-3A	178 × 37 × 22	66.95	0.634	0.556	0.619	0.699
	S-3B	177 × 37 × 24	67.34	0.669	0.650	0.775	0.832
	Total	178 × 37 × 31	65.70	0.535	0.552	0.546	0.662
July-Aug (2019)	S-3A	177 × 37 × 20	57.31	0.507	0.547	0.508	0.646
	S-3B	178 × 34 × 24	53.94	0.462	0.452	0.476	0.565
	Total	178 × 37 × 42	53.38	0.423	0.408	0.387	0.552

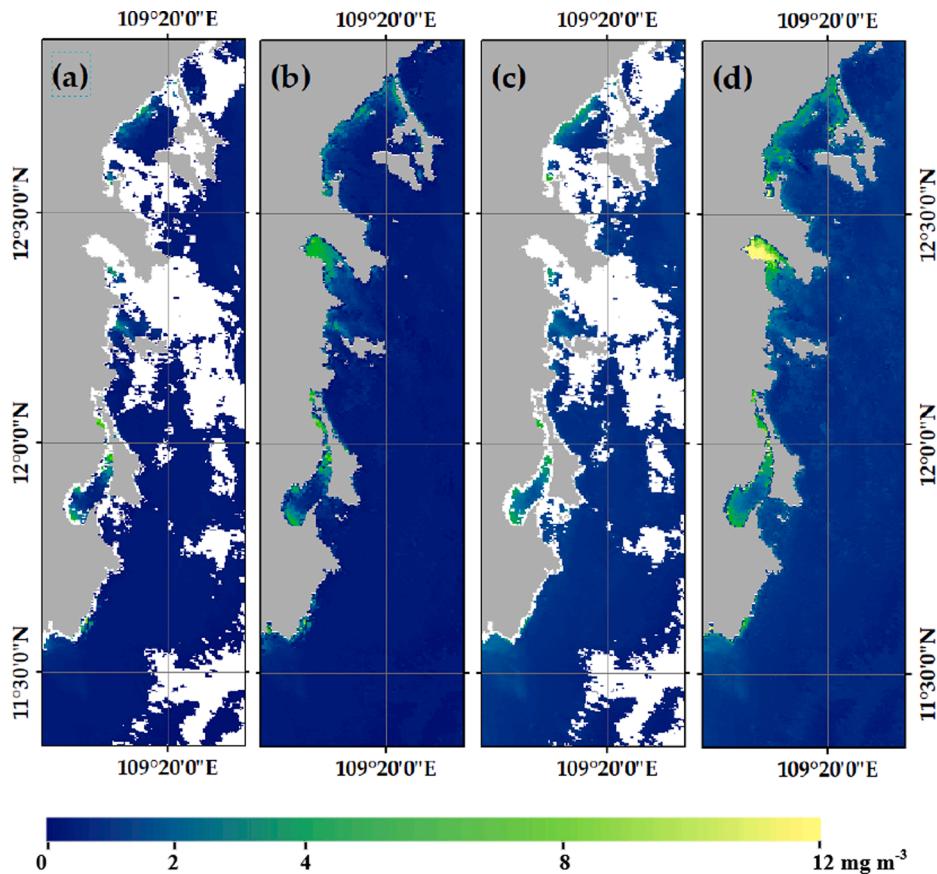


Fig. 2. Illustration of DINEOF performance for the merged datasets in 2019 using OC5-based chl-a products: (a) original C2RCC, (b) filled C2RCC, (c) original DSF, (d) filled DSF.

which yielded the significantly highest errors of MAPE (145.8 %). In general, OC5 seems to be a stable algorithm, with all statistical metrics slightly better in both cases of remotely sensed data S-3A and S-3B.

It is clear that C2RCC performed better while applied in both satellites 3A and 3B (for spectral match-up, see S2 in [supplementary material](#)). Evidence from all errors metrics also confirmed this statement. For example, the RMSE of C2RCC-based products ranges from 0.776 to

1.139 mg m⁻³ while these values are noticeably higher in DSF with 1.451 – 2.087 mg m⁻³. Moreover, MAPE for DSF is significantly higher than C2RCC nearly-three times. The average MAE in C2RCC is 1.6, while for DSF is 2.2. Bias in DSF is around 2.24, while this value improved at least twofold for C2RCC-based products.

Taylor diagrams (Fig. 5) shows a comparison of predicted Chl-a by the different model (for scatter plots, see S3 in [supplementary material](#)).

Table 5

Statistical metrics for the four chl-a algorithms retrieved from each satellite S-3A and S-3B. The atmospheric correction C2RCC (*p*-value < 0.00001).

	R ²	slope	intercept	RMSE (mg m ⁻³)	MAPE (%)	MAE	bias
S-3A							
OC4ME	0.58	0.552	0.494	1.021	41.2	1.693	0.916
OC4	0.57	0.618	0.401	1.037	48.6	1.871	0.807
OC5	0.58	0.586	0.421	1.018	49.4	1.762	0.848
OC6	0.56	0.508	0.503	1.056	35.9	1.657	0.921
S-3B							
OC4ME	0.74	0.728	0.401	0.785	40.5	1.634	1.027
OC4	0.74	0.809	0.377	0.803	38.1	1.650	1.069
OC5	0.75	0.730	0.419	0.776	37.3	1.618	1.060
OC6	0.46	0.496	0.687	1.139	44.0	1.690	1.009

The visualization indicates that OC algorithms applied to S-3A datasets seem to be more stable than S-3B. The diagrams also confirmed the best performance of C2RCC processor. Furthermore, water models OC4ME and OC5 based on C2RCC achieved the best assessment in all cases.

4.3. Synergy of S-3A and S-3B

For the synergy datasets of S-3A and S-3B based on C2RCC atmospheric correction, the statistical results in Fig. 6 show that those three OC algorithms (OC4ME, OC4, OC5) seem to outperform OC6. For instance, RMSE and R², on average, are improved above 20 % and nearly 30 %, respectively. In contrast, only MAPE of OC6 (35.1 %) is lower than three remain algorithms (42.5 – 44.2 %).

Fig. 7 shows the performance of chl-a retrieval based on different OC when processed via DSF using the combination of both satellite 3A and 3B. Across all performance metrics, OC6 resulted in the most accurate estimates (R²: 0.57, RMSE: 1.598 mg m⁻³). OC4ME and OC4 give the same results, with R² being 0.53, and RMSE being 1.872 and 1.411 mg

m⁻³, respectively. OC5 performed poor retrieval in this case, as R² is only 0.46 and RMSE is 1.931 mg m⁻³.

Comparison between both atmospheric corrections show that C2RCC performs significantly better than DSF while considering all four OC algorithms. For instance, the highest RMSE for all C2RCC-based products is only 1.083 mg m⁻³, while the lowest RMSE for DSF products is 1.411 mg m⁻³. In addition, MAPE for all DSF-based products achieved the lowest 87.2 % and highest 152.5 %, while compared to C2RCC, this error only ranged from 35.1 % to 44.2 %. The log transformation of MAE and Bias also confirmed similar results, with those values for C2RCC on average are less than DSF 2–3 times. Overall, the statistics and scatter-plots suggest that the C2RCC processor, OC5 and OC4ME chl-a algorithms can be applied in S-3A or S-3B, as well as the usage of merging of both satellites. Fig. 8 shows the consistency product of chl-a while applying OC5 using the synergy datasets of S-3A and S-3B.

5. Discussion

5.1. Mitigation of the accumulated errors

In order to improve the accuracy of the water model and mitigate the accumulated errors, validation of pre-processing (i.e., atmospheric corrections) is essential before implementing OC algorithms and post-processing (i.e., DINEOF). Unfortunately, the lack of ship-based radiometric measurement leads to the impossible accuracy assessment of the reflectance correction schemes in this research. Notwithstanding AERONET station was successfully installed in Nha Trang Bay since 2011, missing station-based data in the field of remotely sensed ocean color application make it impossible to contribute to the research, as checked in <https://aeronet.gsfc.nasa.gov/>.

On the other hand, we evaluated the retrieval with chl-a field measurements through different ship-based campaigns over three coastal bays (VP, NT, PR) in 2018 and 2019. The results showed that C2RCC outperformed DSF. However, locally C2RCC needs to be applied in order

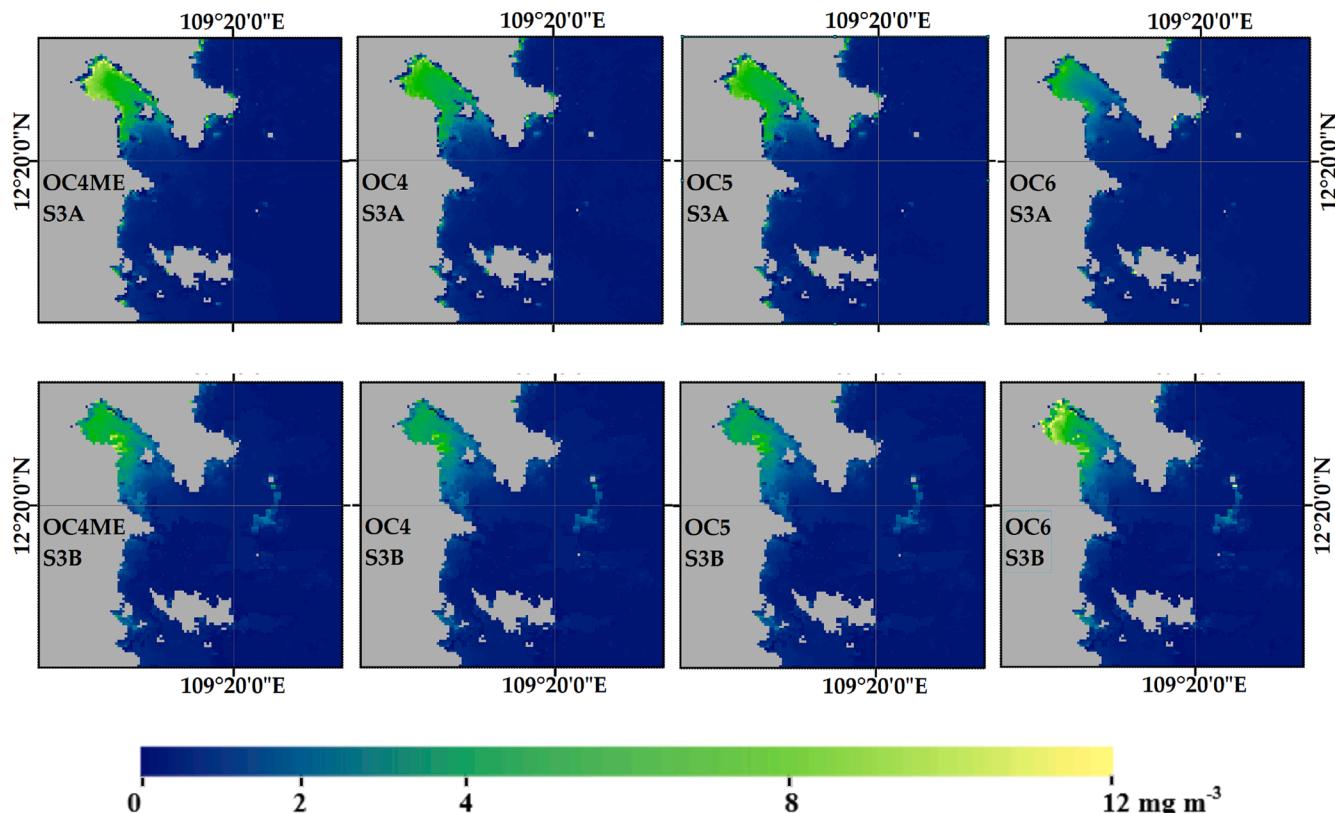


Fig. 3. C2RCC-based Chl-a products over Nha Trang Bay in August 05th, 2019, S-3A (top) and S-3B (bottom).

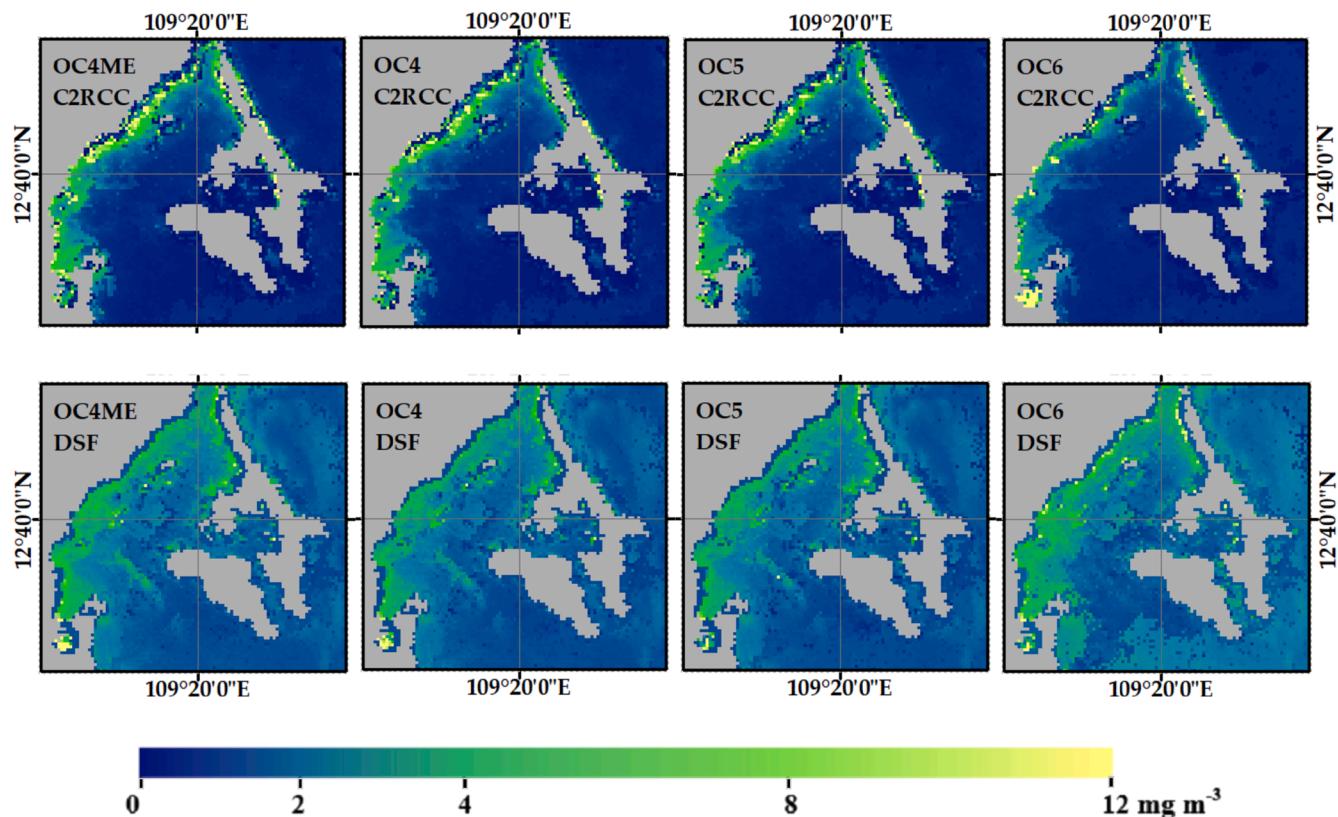


Fig. 4. S-3B-based Chl-a products over Van Phong Bay, October 31st, 2018. Atmospheric correction C2RCC (top) and DSF (bottom).

Table 6

Statistical metrics for the four chl-a algorithms retrieved from each satellite S-3A and S-3B. The atmospheric correction DSF (p -value < 0.00001).

	R ²	slope	intercept	RMSE (mg m ⁻³)	MAPE (%)	MAE	bias
S-3A							
OC4ME	0.49	0.713	1.692	1.755	110.5	2.324	2.246
OC4	0.48	0.624	1.674	1.608	99.1	2.287	2.177
OC5	0.51	0.697	1.531	1.586	98.9	2.156	2.060
OC6	0.58	0.464	1.817	1.451	112.8	2.337	2.223
S-3B							
OC4ME	0.56	0.670	1.825	1.706	130.5	2.422	2.357
OC4	0.56	0.610	1.696	1.526	118.8	2.257	2.185
OC5	0.54	0.601	1.714	1.541	108.5	2.269	2.194
OC6	0.53	0.519	1.997	1.666	145.8	2.528	2.433

to improve the performance of algorithms, as demonstrated in (Kratzer and Plowey, 2021). The new adaptation of DSF for Sentinel 3 integrated into the Acolite was used in this research by comparing with the performance of the machine learning approach C2RCC processor. In DSF, we only applied the threshold of 0.05, which may affect the last results as presented in the validation section. Applying a threshold of 0.03 led to the percentage of missing data in datasets up to nearly 90 %. Consequently, it cannot perform the DINEOF reconstruction with acceptable errors.

In the light of various chl-a retrieval methods, we deployed the original algorithm OC4ME and three new versions 7 of OC4, OC5, OC6. Statistics showed the best performance of OC4ME and OC5 compared to others in all cases. OC4ME was first designed for MERIS, then switching to apply appropriately in Sentinel 3 (Smith et al., 2018) due to the similar orbit characteristics of the instruments (Donlon et al., 2012). The algorithm is based on blue/green band ratio designed for case-1 water or low chl-a concentration (Morel, 2007; Smith et al., 2018). Despite that,

the algorithm was still demonstrated in various coastal areas, continuing to evaluate chl-a algorithms in different water optical properties (Lavigne et al., 2021; Moutzouris-Sidiris and Topouzelis, 2021; Smith et al., 2018; Tilstone et al., 2017). Here, the well performance of OC4ME can be interpretative by the water optical characteristics close to open water, which are also affected by seasonal conditions. The following algorithm, OC5, also worked well except for DSF-based products through merged datasets. Previous works also demonstrated the best performance of OC5 applied to different ocean sensors in numerical coastal areas such as Bay of Bengal and Arabian Sea (SeaWiFS) (Tilstone et al., 2011), Bay of Biscay (MODIS) (Novoa et al., 2012), North West European coastal (MERIS) (Tilstone et al., 2017), Alboran Sea (MODIS) (Gómez-Jakobsen et al., 2016), and Ligurian and North Tyrrhenian Seas (MODIS) (Lapucci et al., 2012). Particularly in the study area, OC5 was demonstrated for the best performance in coastal Vietnam through MERIS (Loisel et al., 2017). With 49 validated samples in this research focus on VP, NT, and PR, further work is essential for confirming reliable ocean color algorithms over larger study areas, as well as OLCI calibration.

5.2. Solution for cloud coverage in tropical areas

In the present study, DINEOF enables a reliable process as already described in the research. The combination of S-3A and S-3B provides the capacity of daily monitoring at full spatial distribution in cloud cover area with acceptable errors of reconstruction processes. In this research, outlier detection in DINEOF was not implemented as we assumed all pixels were corrected after implementing atmospheric correction C2RCC and DSF. It may be affected the errors of reconstruction processes as demonstrated in (Alvera-Azcárate et al., 2012).

5.3. Future implementation

We highlight the benefit of both satellite S-3A and S-3B by applying multi-processing steps to retrieve chl-a. The research opens new avenues

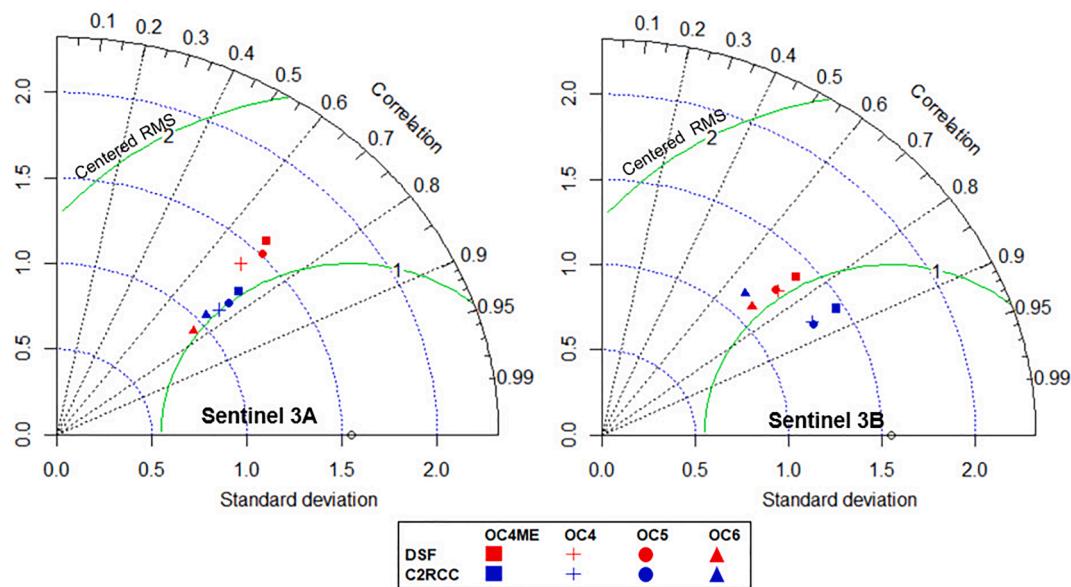


Fig. 5. Taylor diagram represent the performance of different chl-a algorithms based on C2RCC and DSF, S-3A (left) and S-3B (right).

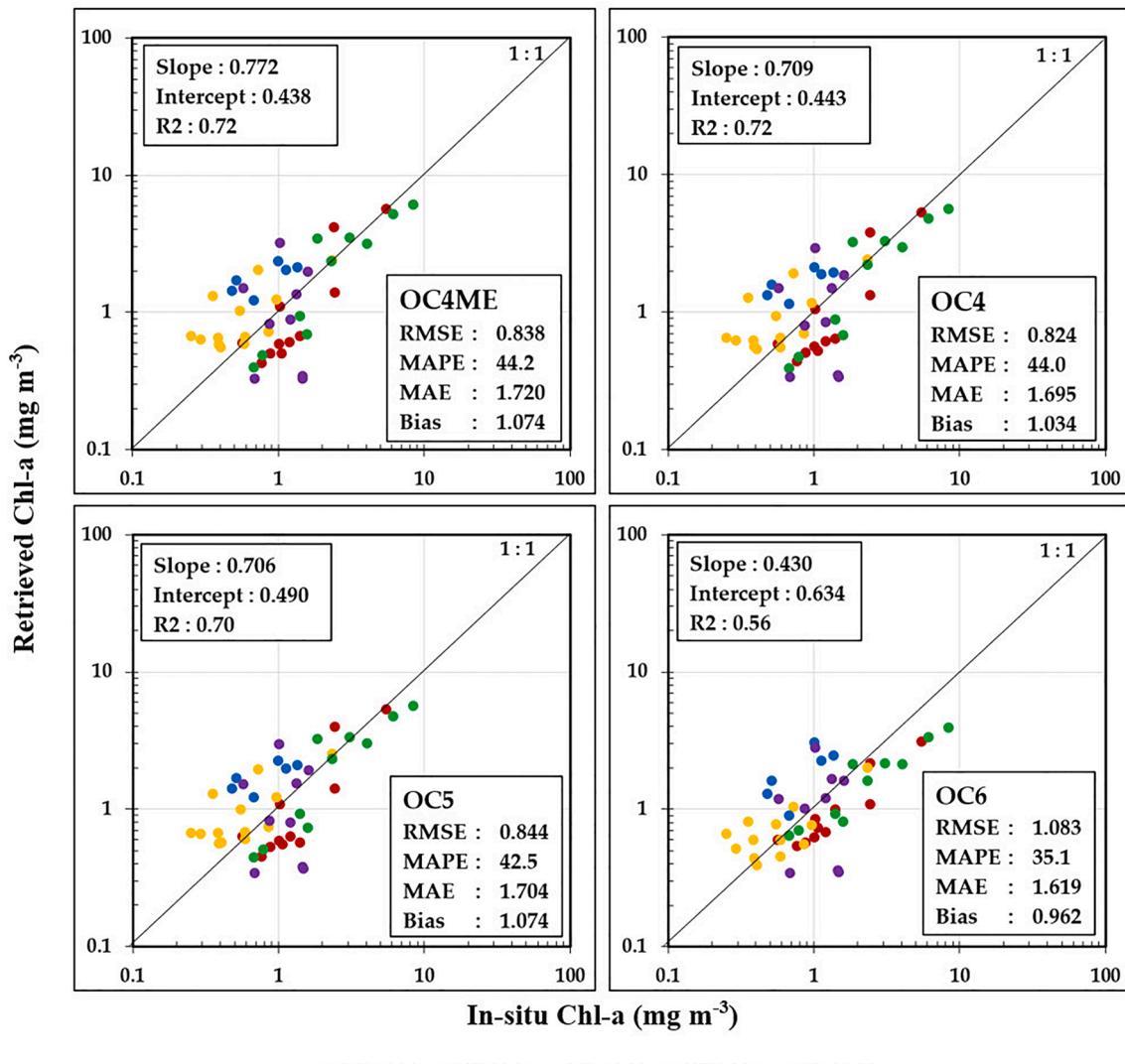


Fig. 6. Performance of C2RCC-based chl-a products derived from different algorithms when combining S-3A and S-3B. (p -value < 0.00001).

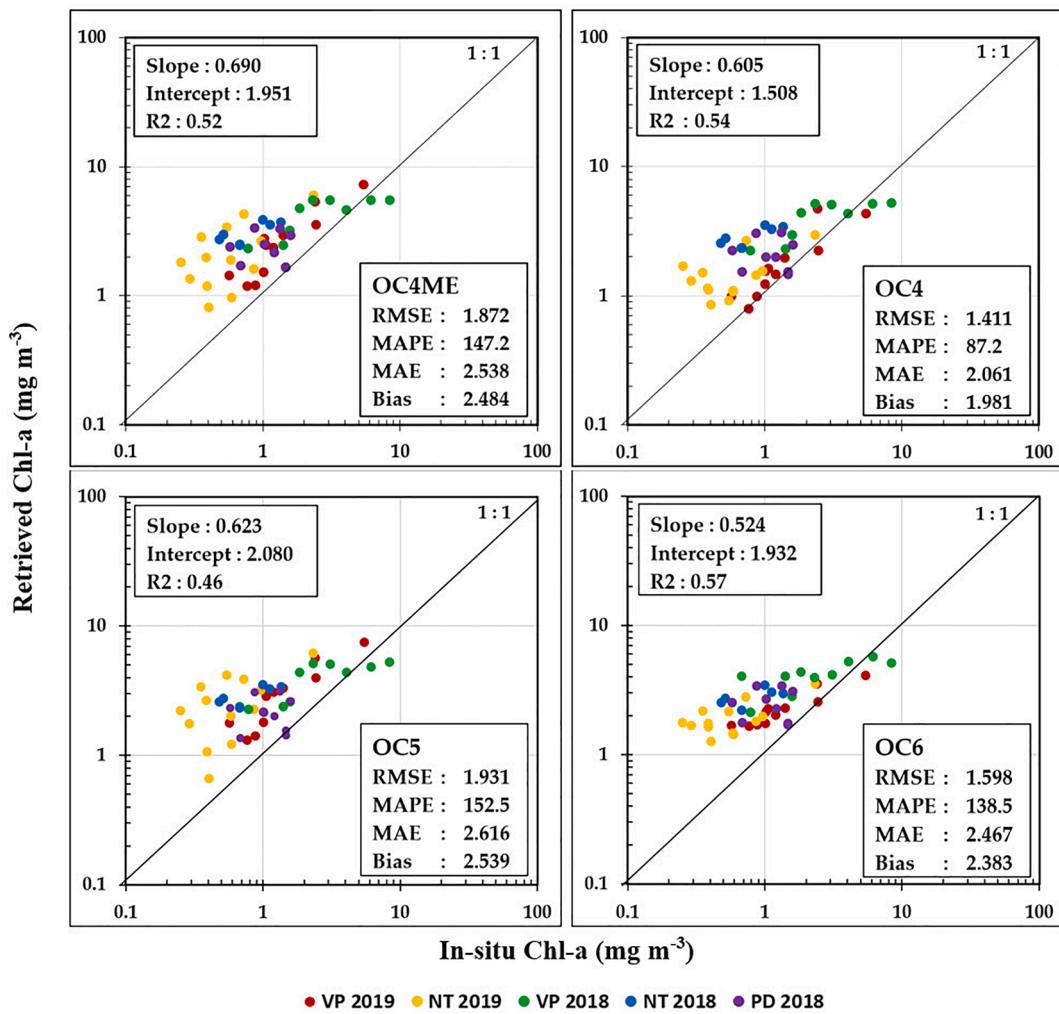


Fig. 7. Performance of DSF-based chl-a products derived from different algorithms when combining S-3A and S-3B. (p -value < 0.00001).

for future work to include other biogeochemical variables of coastal ecosystems (i.e., turbidity, suspended particulate matter) toward a profound understanding of remote sensing ocean color in future satellite generations and focusing on coastal Vietnam. Multi pre- and post-processing steps for the quantitative estimation maybe lead to accumulated errors. Therefore, further independent evaluation of each step (atmospheric correction processors and bio-optical algorithms) is essential to optimize the most suited approach.

Our study covered short periods as used only for examination purposes. Exploiting time series of satellite remote sensed data would allow the temporal monitoring over the coastal areas. Consequently, this can help improve the knowledge of seasonal characteristics, trends, and the impact of human activities. Cloud cover is a major issue when capturing optical images in the tropical monsoon climate areas. Therefore, gap-filling tools and the combination of different optical instruments are essential for improving the capacity of optimization spatio-temporal monitoring.

6. Conclusion

Our study demonstrated the statistical performance for predicting chl-a over small water areas in the Vietnam southern coastal against in-situ measurements in different locations and periods. Firstly, atmospheric correction C2RCC has been confirmed to work well in the research area. Secondly, chl-a algorithm OC4ME and OC5 based on C2RCC processors can be applied for chl-a retrieval with acceptable

errors. Finally, DINEOF performed reasonably well in both cross-validation and in-situ assessment for datasets with limited temporal dimensions. The study showed the promising solution for producing chl-a while applying in cloud areas through multi-processing steps. The capacity of each satellite Sentinel 3A/B, as well as the synergy datasets, emphasize the high frequency temporal monitoring in the coastal regions. The research contributed to the field of remotely sensed ocean color in coastal Vietnam, enabling the potential of quantitative estimation using different bio-optical characteristics following OLCI sensor and future generations.

CRediT authorship contribution statement

Nguyen An Binh: Conceptualization, Methodology, Software, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Pham Viet Hoa:** Conceptualization, Supervision, Resources, Writing – review & editing. **Giang Thi Phuong Thao:** Methodology, Software, Validation, Data curation, Writing – review & editing. **Ho Dinh Duan:** Methodology, Supervision, Resources, Writing – review & editing. **Phan Minh Thu:** Methodology, Investigation, Resources, Formal analysis, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial

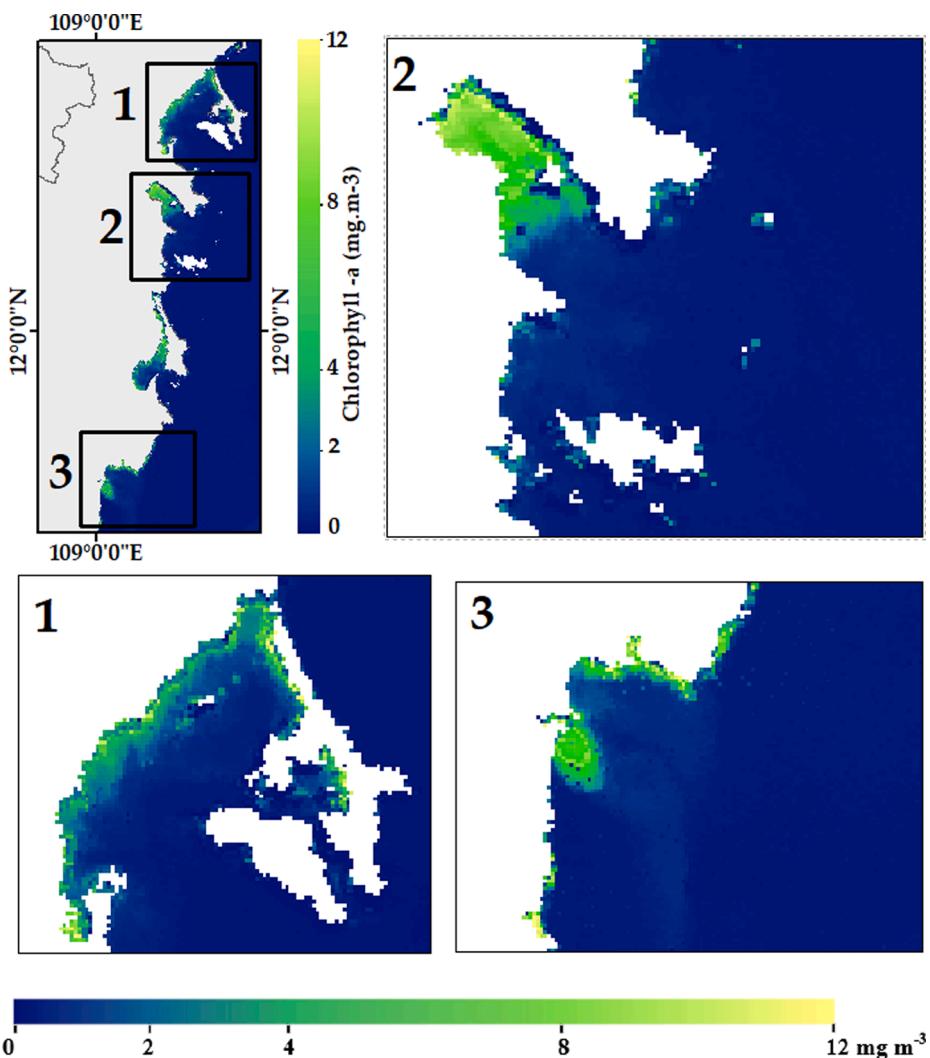


Fig. 8. Synergy of both satellite S-3A and S-3B for Chl-a derived from OC5, July 31st, 2019, Atmospheric correction C2RCC.

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the project CSCL21.01/22-23. We thank Vietnam Academy of Science and Technology (VAST) for funding the research project grant number CSCL21.01/22-23. For field data sharing, we thank national projects of VT-UD.12/18-20, ĐT-2018-40502-ĐL2.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jag.2022.102951>.

References

- Alvera-Azcárate, A., Barth, A., Rixen, M., Beckers, J.M., 2005. Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: Application to the Adriatic Sea surface temperature. *Ocean Model.* 9 (4), 325–346.
- Alvera-Azcárate, A., Sirjacob, D., Barth, A., Beckers, J.-M., 2012. Outlier detection in satellite data using spatial coherence. *Remote Sens. Environ.* 119, 84–91.
- Alvera-Azcárate, A., Vanhellemont, Q., Ruddick, K., Barth, A., Beckers, J.-M., 2015. Analysis of high frequency geostationary ocean colour data using DINEOF. *Estuar. Coast. Shelf Sci.* 159, 28–36.
- Alvera-Azcárate, A., Barth, A., Parard, G., Beckers, J.-M., 2016. Analysis of SMOS sea surface salinity data using DINEOF. *Remote Sens. Environ.* 180, 137–145.
- Anderson, C.R., Kudela, R.M., Kahru, M., Chao, Y.i., Rosenfeld, L.K., Bahr, F.L., Anderson, D.M., Norris, T.A., 2016. Initial skill assessment of the California Harmful Algae Risk Mapping (C-HARM) system. *Harmful Algae* 59, 1–18.
- Barth, A., Alvera-Azcárate, A., Troupin, C., Beckers, J.-M., 2022. DINCAE 2.0: multivariate convolutional neural network with error estimates to reconstruct sea surface temperature satellite and altimetry observations. *Geosci. Model Dev.* 15 (5), 2183–2196.
- Beckers, J.M., Rixen, M., 2003. EOF calculations and data filling from incomplete oceanographic datasets. *J. Atmos. Ocean. Technol.* 20. doi: 10.1175/1520-0426(2003)020<1839:ECADFF>2.0.CO;2.
- Brockmann, C., Doerffer, R., Peters, M., Stelzer, K., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. In: European Space Agency, (Special Publication) ESA SP.
- Doerffer, R., Schiller, H., 2007. The MERIS case 2 water algorithm. *Int. J. Remote Sens.* 28 (3–4), 517–535.
- Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sens. Environ.* 120, 37–57.
- Falkowski, P., 2012. Ocean science: the power of plankton. *Nature* 483 (7387), S17–S20.
- Gohin, F., Druon, J.N., Lampert, L., 2002. A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. *Int. J. Remote Sens.* 23, 1639–1661. <https://doi.org/10.1080/01431160110071879>.
- Gómez-Jakobsen, F., Mercado, J.M., Cortés, D., Ramírez, T., Salles, S., Yebra, L., 2016. A new regional algorithm for estimating chlorophyll-a in the Alboran Sea (Mediterranean Sea) from MODIS-Aqua satellite imagery. *Int. J. Remote Sens.* 37 (6), 1431–1444.
- Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Appl. Opt.* 33 (3), 443.

- Gordon, H.R., Clark, D.K., Brown, J.W., Brown, O.B., Evans, R.H., Broenkow, W.W., 1983. Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates. *Appl. Opt.* 22 (1), 20.
- Ha, N., Koike, K., Nhuan, M., 2013. Improved accuracy of chlorophyll-a concentration estimates from MODIS Imagery using a two-band ratio algorithm and geostatistics: as applied to the monitoring of eutrophication processes over Tien Yen Bay (Northern Vietnam). *Remote Sens.* 6 (1), 421–442.
- Han, Z., He, Y., Liu, G., Perrie, W., 2020. Application of DINCAE to reconstruct the gaps in chlorophyll-a satellite observations in the South China sea and West Philippine Sea. *Remote Sens.* 12 (3), 480.
- IOCCG, 2021. IOCCG Report Number 20, Observation of Harmful Algal Blooms with Ocean Colour Radiometry, in: Reports and Monographs of the International Ocean Colour Coordinating Group.
- Jeffrey, S.W., Mantoura, R.F.C., Wright, S.W., 1997. Spectrophotometric and fluorometric equations in common use in oceanography. *Phytoplankt. Pigment. Oceanogr. Guidel. to Mod. Methods* 48.
- Ji, C., Zhang, Y., Cheng, Q., Tsou, JinYeu, Jiang, T., Liang, X.S., 2018. Evaluating the impact of sea surface temperature (SST) on spatial distribution of chlorophyll-a concentration in the East China Sea. *Int. J. Appl. Earth Obs. Geoinf.* 68, 252–261.
- Ji, C., Zhang, Y., Cheng, Q., Tsou, J.Y., 2021. Investigating ocean surface responses to typhoons using reconstructed satellite data. *Int. J. Appl. Earth Obs. Geoinf.* 103, 102474.
- Kitsiou, D., Karydis, M., 2011. Coastal marine eutrophication assessment: a review on data analysis. *Environ. Int.* 37 (4), 778–801.
- Konik, M., Kowalewski, M., Bradtke, K., Darecki, M., 2019. The operational method of filling information gaps in satellite imagery using numerical models. *Int. J. Appl. Earth Obs. Geoinf.* 75, 68–82.
- Kratzer, S., Plowey, M., 2021. Integrating mooring and ship-based data for improved validation of OLCI chlorophyll-a products in the Baltic Sea. *Int. J. Appl. Earth Obs. Geoinf.* 94, 102212.
- Lapucci, C., Ampolo Rella, M., Brandini, C., Ganzin, N., Gozzini, B., Maselli, F., Massi, L., Nuccio, C., Ortolani, A., Trees, C., 2012. Evaluation of empirical and semi-analytical chlorophyll algorithms in the Ligurian and North Tyrrhenian Seas. *J. Appl. Remote Sens.* 6 (1), 063565-1.
- Lavigne, H., Van der Zande, D., Ruddick, K., Cardoso Dos Santos, J.F., Gohin, F., Brotas, V., Kratzer, S., 2021. Quality-control tests for OC4, OC5 and NIR-red satellite chlorophyll-a algorithms applied to coastal waters. *Remote Sens. Environ.* 255, 112237.
- Li, Y., He, R., 2014. Spatial and temporal variability of SST and ocean color in the gulf of maine based on cloud-free SST and chlorophyll reconstructions in 2003–2012. *Remote Sens. Environ.* 144, 98–108.
- Loisel, H., Mangin, A., Vantrepotte, V., Dessailly, D., Ngoc Dinh, D., Garnesson, P., Ouillon, S., Lefebvre, J.-P., Mériaux, X., Minh Phan, T., 2014. Variability of suspended particulate matter concentration in coastal waters under the Mekong's influence from ocean color (MERIS) remote sensing over the last decade. *Remote Sens. Environ.* 150, 218–230.
- Loisel, H., Vantrepotte, V., Ouillon, S., Ngoc, D.D., Herrmann, M., Tran, V., Mériaux, X., Dessailly, D., Jamet, C., Duhaut, T., Nguyen, H.H., Van Nguyen, T., 2017. Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS ocean color sensor (2002–2012). *Remote Sens. Environ.* 190, 217–232.
- Mercado, J.M., Gómez-Jakobsen, F., Cortés, D., Yebra, L., Salles, S., León, P., Putzeys, S., 2016. A method based on satellite imagery to identify spatial units for eutrophication management. *Remote Sens. Environ.* 186, 123–134.
- Morel, A.D.A., 2007. Pigment index retrieval in Case 1 waters. Algorithm Theor. basis Doc. 2.9.
- Moutzouris-Sidiris, I., Topouzelis, K., 2021. Assessment of Chlorophyll-a concentration from Sentinel-3 satellite images at the Mediterranean Sea using CMEMS open source in situ data. *Open Geosci.* 13. doi: 10.1515/geo-2020-0204.
- Nechad, B., Alvera-Azcarate, A., Ruddick, K., Greenwood, N., 2011. Reconstruction of MODIS total suspended matter time series maps by DINEOF and validation with autonomous platform data. *Ocean Dyn.* 61 (8), 1205–1214.
- Ngoc, D.D., Loisel, H., Jamet, C., Vantrepotte, V., Duforet-Gaurier, L., Minh, C.D., Mangin, A., 2019. Coastal and inland water pixels extraction algorithm (WiPE) from spectral shape analysis and HSV transformation applied to Landsat 8 OLI and Sentinel-2 MSI. *Remote Sens. Environ.* 223, 208–228.
- Ngoc, D.D., Loisel, H., Vantrepotte, V., Chu Xuan, H., Nguyen Minh, N., Verpoorter, C., Meriaux, X., Pham Thi Minh, H., Le Thi, H., Le Vu Hong, H., Nguyen Van, T., 2020. A simple empirical band-ratio algorithm to assess suspended particulate matter from remote sensing over coastal and inland waters of vietnam: application to the VNREDSat-1/NAOMI sensor. *Water (Switzerland)* 12 (9), 2636.
- Novoa, S., Chust, G., Sagarnaga, Y., Revilla, M., Borja, A., Franco, J., 2012. Water quality assessment using satellite-derived chlorophyll-a within the European directives, in the southeastern Bay of Biscay. *Mar. Pollut. Bull.* 64 (4), 739–750.
- O'Reilly, J., Maritorena, S., 2000. Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3.
- O'Reilly, J.E., Werdell, P.J., 2019. Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6. *Remote Sens. Environ.* 229 <https://doi.org/10.1016/j.rse.2019.04.021>.
- Oke, P.R., Brasington, G.B., Griffin, D.A., Schiller, A., 2010. Ocean data assimilation: a case for ensemble optimal interpolation. *Aust. Meteorol. Oceanogr. J.* 59. doi: 10.22499/2.5901.008.
- O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res. Ocean.* 103 (C11), 24937–24953.
- Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., Alikas, K., Kangro, K., Gurlin, D., Hà, N., Matsushita, B., Moses, W., Greb, S., Lehmann, M.K., Ondrusk, M., Oppelt, N., Stumpf, R., 2020. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sens. Environ.* 240, 111604.
- 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 (6), 7404.
- Smith, M.E., Robertson Lain, L., Bernard, S., 2018. An optimized Chlorophyll a switching algorithm for MERIS and OLCI in phytoplankton-dominated waters. *Remote Sens. Environ.* 215, 217–227.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* 106 (D7), 7183–7192.
- Tilstone, G.H., Angel-Benavides, I.M., Pradhan, Y., Shutler, J.D., Groom, S., Sathyendranath, S., 2011. An assessment of chlorophyll-a algorithms available for SeaWiFS in coastal and open areas of the Bay of Bengal and Arabian Sea. *Remote Sens. Environ.* 115 (9), 2277–2291.
- Tilstone, G., Mallor-Hoya, S., Gohin, F., Couto, A.B., Sá, C., Goela, P., Cristina, S., Airs, R., Icely, J., Zühlke, M., Groom, S., 2017. Which ocean colour algorithm for MERIS in North West European waters? *Remote Sens. Environ.* 189, 132–151.
- Tilstone, G.H., Pardo, S., Dall'Olmo, G., Brewin, R.J.W., Nencioni, F., Dessailly, D., Kwiatkowska, E., Casal, T., Donlon, C., 2021. Performance of Ocean Colour Chlorophyll a algorithms for Sentinel-3 OLCI, MODIS-Aqua and Suomi-VIIRS in open-ocean waters of the Atlantic. *Remote Sens. Environ.* 260 <https://doi.org/10.1016/j.rse.2021.112444>.
- Vanhellemont, Q., Ruddick, K., 2021. Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. *Remote Sens. Environ.* 256 <https://doi.org/10.1016/j.rse.2021.112284>.
- Werdell, P.J., Bailey, S.W., 2005. An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ.* 98 (1), 122–140.