Evaluation of Satellite Sensors to Compute Chlorophyll-a Concentration in the Northeastern Arabian Sea: A Validation Approach

Vinaya Kumar Vase1,2 • Nakhawa Ajay2 • Rajan Kumar2 • Jayasankar Jayaraman2 • Prathibha Rohit1,2

Received: 14 July 2021 / Accepted: 2 August 2022 / Published online: 17 August 2022 © Indian Society of Remote Sensing 2022

Abstract

The primary productivity of an aquatic system like the Arabian Sea is majorly supported by the concentration of Chlorophyll-a (Chl-a/C_a) pigment. The present study is narrated to validate the Chl-a data set retrieved from the prominent ocean color sensors (OC3M-MODIS, OC-OCM2, and OC3V-VIIRS) with sea truth data, collected from 204 stations for three years (2015–2017). The in situ concentrations of Chl-a depict the geographic region under the mesotrophic and eutrophic spans with a mean of 1.36 mg m^{-3} (0.1 > C_a > 1.0 mg m^{-3}). The ratio of C_a^{OCM2}/C_a^{in-situ} was 0.97 ± 0.27 mg m^{-3} (n = 199), but the ratios were higher with C_a^{VIIRS}/C_a^{in-situ} is 1.75 ± 0.79 mg m^{-3} (n = 170) and C_a^{MODIS}/C_a^{in-situ} is 2.53 ± 1.42 mg m^{-3} (n = 158). The coefficient of determination proclaims a significant relationship for MODIS (R^2 = 0.36; p < 0.001), followed by OCM2 (R^2 = 0.32; p < 0.001) and VIIRS (R^2 = 0.19; p < 0.001). The OCM2 showed the lowest RMSE at 0.13, which is relatively lower than the reference error limit by global ocean color missions at 0.35. In overall performance among the three algorithms evaluated for the region, the OCM2 will provide a better estimation of Chl-a with a prediction of 32% accuracy and 34.3% of bias. The log bias values for MODIS (0.35) and VIIRS (0.20) algorithms indicate the overestimation of Chl-a with in situ Chl-a, but the OCM2 algorithm is suitable in the region with a negligible bias (-0.03). The biogeochemical processes and ecosystem characteristics are dynamic from region to region, as yet in its urgent need to validate global sensors to fine-tune the regional algorithms periodically.

Keywords Sensors • Productivity pigment • Validation • Eutrophic • Northern Arabian Sea

Key Points

• The study generates in situ Chlorophyll-a (Chl-a) for the northern Arabian Sea, which is a key productivity parameter to understand the dynamics of the ecosystem
• The measured mean Chl-a was estimated at 1.36 mg m^{-3}, which indicated that the study region is categorized under the eutrophic status.
• Overestimation (RMSE and Bias) of Chl-a was observed for the global sensors like MODIS and VIIRS.
• The OCM2 sensor performed better with 34.4% bias and 32% variance than the other two sensors, i.e., MODIS and VIIRS sensors in the region

Introduction

The major driving force to restrain the earth’s climate is the ocean, but the extent of interaction is limited in understanding. The major bottlenecks to understand the biogeochemical process due to limited onboard cruises and other available ocean-based programs. The researchers would get a synoptic view of physicochemical and biological changes with the higher resolutions of spatial and temporal scales from the data collected from satellite-borne ocean color sensors (Werdell et al., 2003). A precise assessment of phytoplankton biomass in the oceans is required to get an insight into the ocean carbon cycle. A large number of
in situ data observations are required to validate the satellite products to build up algorithms (Lavigne et al., 2012). With the advancement of satellite technology and the development of precise sensors, viz., MODIS (Moderate Resolution Imaging Spectroradiometer; Esaias et al., 1998), OCTS (Ocean Color and Temperature Sensor, Tanii et al., 1991), MERIS (Medium Resolution Imaging Spectrometer, Rast & Bezy, 1995), and SeaWiFS (Sea-viewing Wide Field-of-view Sensor, McClain et al., 1998a, 1998b), the research investigations have been considered, especially in the coastal and marine research. The ability of remote sensing technology to cover wider spatiotemporal realms has made them the method of choice in marine ecological study. The color of the water of an aquatic system is a function of the concentration of components like organic carbon, other particulate matter, and pigments from phytoplankton (Morel & Prieur, 1977). In situ data as ground truth can be used to validate and calibrate the satellite products and one such effort was attempted by the NASA Ocean Biology Processing Group (OBPG), in addition, maintained a repository of bio-optical data (Werdell & Bailey, 2005).

In the Earth’s Carbon cycle and budget, the oceanic system plays a very prominent role to decode the same need to have a fair understanding of the magnm and primary production rate on a global scale and similarly in diverse and unique sub-systems of the oceans (Prasad & Haedrich, 1994). In aquatic systems, Chlorophyll-a concentration is most often used as a proxy for primary productivity and the difference over a wider spatial and temporal scale becomes important to understanding the carbon budget at a global scale (Platt et al., 1991; Sarmento et al., 2004; Tomlinson et al., 2004). In the global carbon cycle, the phytoplankton plays a prominent role; it is high time to estimate the spatial and temporal distribution and variability of Chl-a for modern oceanography (Claustre et al., 2010). Although Chl-a is the most abundant biological oceanic measurement, however, the data generated is scarce in comparison with the number of other physical observations studied (e.g., temperature and salinity). Research investigations have been made on the empirical relationship between satellite-derived ocean color and in situ Chl-a in recent years (Clark, 1981; Shang et al., 2014; Tilstone et al., 2013).

The regional field data (ground truths) are limited from single to few field trips is often a limiting factor to validating satellite-derived Chl-a (Barbini et al., 2005). To validate the existing algorithms and newer ones to be developed, the high-quality in situ data over a larger temporal and spatial scale is a must to attain the objective. The Arabian Sea marine ecosystem is unique in more than one way; it is under marked continental influence and has a profound effect on the monsoon, which propels perennial productivity by ensuring vertical mixing of nutrients (Banse & English, 2000). The ground truth sampling is consorting with many limitations like time management, huge cost, sampling errors, disparities in sampling protocols, spatial coverage the paucity of wide-ranging observations, etc. The sensors developed for the ocean color products (SeaWiFS, MODIS Terra, MODIS Aqua, OCM2, and VIIRS) with the associated limitations will provide a synoptic view of the globe. There were very scanty in situ data sets accessible and few validation trials were attempted for available ocean color products before the present study region (Al-Naimi et al., 2017; Raman, 2013). The key objective of the present research attempt is to generate in situ Chl-a data, which could characterize the regional productivity of the ecosystem and its dynamics. In addition, the study evaluates the performance of ocean color satellite sensors (MODIS, OCM2, and VIIRS) with in situ Chl-a data to suggest a suitable sensor database for large spatial scale application in the northeastern Arabian Sea.

Materials and Methods

Sampling Design

To generate robust and accurate algorithms, in addition, to validating satellite-derived Chl-a concentration at a regional scale will be limited because of the insufficiency of in situ datasets. The study area comes under the northeastern Arabian Sea and is located on the northwest coast of India’s Exclusive Economic Zone (EEZ). The in situ samplings were carried out at two locations, i.e., off Veraval and Okha coasts, Gujarat (Arabian Sea; Fig. 1). Onboard scientific sampling cruises were conducted at monthly temporal intervals for a period of three years (2015 to 2017). The in situ samples were collected randomly from the 204 sampling stations during the study period, but the sample size was less during the monsoon season (June–September) due to the erratic sea conditions. To understand the temporal variations in Chl-a, the data were categorized into four main seasons, i.e., winter monsoon (December–March), spring inter-monsoon (April–May), monsoon (June to September), and fall inter-monsoon (October–November) (Raman, 2013).

Estimation of In Situ Chlorophyll-a Concentration

The most practiced conventional approaches to measure Chl-a are fluorometry (Holm-Hansen et al., 1965), spectrophotometry (Lorenzen, 1967), and chromatography (Mantoura & Llewellyn, 1983) with different levels of accuracy and precision. Among three approaches, the most accurate method was high-performance liquid
chromatography (HPLC, Hooker, et al., 2009), which provides the concentrations of a large spectrum of phytoplankton accessory pigments in addition to Chl-a. The fluorescence method is the only one that has not been included in the global reanalysis of Chl-a, for example, open ocean climatologies are among the three main approach methodologies that exist for measuring Chl-a (i.e., water sampling, ocean color, and induced fluorescence), (Gregg & Conkright, 2001). The spectroscopic approach was adopted to evaluate the performance of the Chl-a value of different satellite sensors because we aimed principally to analyze Chl-a, not for any pigments.

The surface water samples (0.5 L for each) were collected at each sampling station. After collection, water samples were protected from sunlight and stored in chilled water to arrest the degradation of Chlorophyll-a pigments on-board. The collected samples were properly labeled on board and kept in the icebox with gel ice in the dark and cold region of the boat. The sample holding time was about four hours before reaching the laboratory for filtration and further analysis. Samples were filtered through 47 mm GF/F filter paper using a vacuum filter for immediate filtration to avoid time lag. Followed, place the filter paper in a 90% acetone solution; later it was kept in the dark for 24 h in a refrigerator. The extracted material was separated using a centrifuge (R 24 Research Centrifuge, REMI, India) and the absorbance concentration was calculated by using a spectrophotometer at different wavelengths of 750 nm, 665 nm, 645 nm, and 630 nm (Vase et al., 2018). Decant the supernatant into a 10 mm path length cuvette and measure the extinction at different wavelengths against a cell containing 90% acetone. The amount of Chlorophyll-a in the sample was computed using the formula (Strickland & Parsons, 1972):

$$C(\text{Chlorophyll-a}) = 11.6E665 - 1.31E645 - 0.14E630$$

where $E$ is the absorbance at different wavelengths above and corrected by the 750 nm, and C is the amount of Chlorophyll-a estimated.

**Satellite/In Situ Matchups**

To evaluate the satellite Chlorophyll-a of selected sensors with in situ data, we have collected all images of
the Moderate Resolution Imaging Spectroradiometer (MODIS) of NASA, Visible Infrared Imaging Radiometer Suite (VIIRS) of NASA, and Ocean Color Monitor onboard OCEANSAT-2 of ISRO (OCM2). The mean bias (MB) and mean absolute error (MAE) of Chl-a retrieval were estimated for the MODIS (1.183 MB and 1.676 MAE) and VIIRS (1.043 MB and 1.482 MAE) algorithms (Seegers et al., 2018). The variable range and target error budget was estimated for the OCM2 algorithm at 0.05 to 30 mg m$^{-3}$ and < 30%, respectively (Navalgund, 2008). Out of a total in situ data points, the obtained matchups for the satellite products are, i.e., MODIS ($n = 158$), VIIRS ($n = 170$), and OCM2 ($n = 199$) due to the sun glint, temporal mismatch of satellite coverage, cloudy/haze, and monsoonal conditions during the sampling period. The 8-day composite L3 Chl-a data product with a high resolution of $\sim$ 4 km$^2$/pixel was retrieved to evaluate the variability in the sensors dataset in the region. A stringent comparison between satellite and in situ data should limit the time difference between the two measurements to within $\pm$ 2–3 h. The in situ Chl-a datasets were a matchup in time (8-day temporal composite matchup) and spatial (latitude and longitude) with the satellite-derived dataset (Al-Naimi et al., 2017). The 8-day composite products were preferred for the region instead of daily satellite data, while the temporal threshold may not decrease the accuracy (Johnson et al., 2013).

**Statistical Analysis**

The in situ and satellite Chl-a concentrations were compared to understand the disparities/similarities in variability, correlation, accuracy, and error. The performance indicators like a satellite to in situ ratios, semi-interquartile range (SIQR), median satellite to in situ ratio, and median absolute percentage difference (MPD) were estimated to understand the overestimation/underestimation of satellite Chl-a. The indication of the spread of the data and measurement of uncertainty is explained by SIQR in addition to how accurately the satellite Chl-a agrees with in situ measurements.

\[
\text{SIQR} = \frac{Q_3 - Q_1}{2}
\]

where $Q_1$ is the 25th percentile, and $Q_3$ is the 75th percentile ($Q_2$ would be the median value). The median ratio indicates an overall bias. The MPD (median absolute percentage difference) is calculated as the median of the individual absolute percentage differences (Bailey & Werdell, 2006):

\[
PD_i = 100 \times \frac{|X_i - Y_i|}{Y_i}
\]

where $X$ is the satellite value, and $Y$ is the in situ value.

The root mean square (RMS) and the mean difference (bias) in percentage are used to assess the differences between the in situ and satellite-derived data (Marrari et al., 2006):

\[
\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)^2} \times 100
\]

\[
\text{bias} = \bar{x} = \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right) \times 100
\]

\[
x = \frac{S - I}{I}
\]

where $S$ is the satellite data, $I$ is in situ data, and $n$ is the number of data pairs. For a normally distributed $x$, RMS should equal the standard deviation. Further, because the natural distribution of Chl-a is lognormal (Campbell, 1995), error estimates were also made on the logarithmically transformed (base 10) data:

\[
\log_{\text{RMS}} = \sqrt{\frac{\sum_{i=1}^{n} (\log(S) - \log(I))^2}{n}}
\]

\[
\log_{\text{bias}} = \frac{\sum_{i=1}^{n} \log(S) - \log(I)}{n}
\]

Generally, error estimates cannot be expressed as percentages because they are transformed logarithmically. These error estimates have been used in recent publications to validate SeaWiFS global and regional estimates of Chl-a (Zhang et al., 2006) and to describe the performance of the ocean color algorithms (O’Reilly et al., 2000).

**Results**

The significant temporal variations (monthly) were shown in Fig. 2, illustrating the maximum during the winter monsoon (1.52 ± 0.48 mg m$^{-3}$) and minimum during the monsoon season (1.25 ± 0.50 mg m$^{-3}$) ($p < 0.05$). The representation of log Chl-a data in the histogram plot depicts the study region was categorized under the mesotrophic and eutrophic categories (Chl-a concentration of $> 0.1$ mg m$^{-3}$) (Fig. 3). The median of Chl-a value was highest for the MODIS at 2.92 mg m$^{-3}$ (ranged from 0.28 to 18.75 mg m$^{-3}$), followed by VIIRS as 2.43 mg m$^{-3}$ (ranged from 0.36 to 11.82 mg m$^{-3}$), OCM2 as 1.35 mg m$^{-3}$ (ranged from 0.57 to 2.38 mg m$^{-3}$) and in situ as 1.36 mg m$^{-3}$ (ranged from 0.50–2.96 mg m$^{-3}$) (Table 1; Fig. 4).

The mean ratio of $C_a^{\text{OCM2}}/C_a^{\text{In-situ}}$ is closer to 1 with a narrow deviation (i.e., 0.97 ± 0.27 mg m$^{-3}$), in contrast to the higher ratio for $C_{\text{VIIRS}}^{\text{OCM2}}/C_a^{\text{In-situ}}$ (1.75 ± 0.79 mg m$^{-3}$)
and $C_a^{\text{MODIS}}/C_a^{\text{In-situ}}$ (2.53 ± 1.42 mg m$^{-3}$), which depicts the overestimation of Chl-a for the later satellite dataset (Fig. 5). The correlation analysis between the in situ and satellite Chl-a data set revealed a better correlation with MODIS (0.602, $p < 0.001$), followed by VIIRS (0.589, $p < 0.001$), and OCM2 (0.562, $p < 0.001$) (Fig. 6). The OCM2 satellite data product is having less
uncertainty in comparison with the other two datasets in the region.

The SIQR is an indication of the uncertainty of the satellite data in comparison with the in situ Chl-α value, and the lower uncertainty was observed in Ca_OCM2 versus Ca_In-situ (0.27 mg m$^{-3}$), followed by Ca_VIIRS versus Ca_In-situ (0.79 mg m$^{-3}$) and Ca_MODIS versus Ca_In-situ (1.42 mg m$^{-3}$). The Ca_OCM2 is scattering nearly around the 1:1 regression line, but a scatter pattern was observed with other datasets Ca_VIIRS/Ca_In-situ and Ca_MODIS/Ca_In-situ. The mean absolute percentage differences (MAPD) estimated for three different sensors were shown in Table 2. The lower MAPD value was estimated for the OCM2 (35.01%), which is having better accuracy with the in situ data than the other sensors, i.e., VIIRS (142.94%) and MODIS (192.39%). The bias percentage is lower in the case of OCM2 at 34.37%, followed by VIIRS (154.88%) and MODIS (294.02%). The RMS error was observed lower for the OCM2 (49.01%), followed by VIIRS (212.72%) and MODIS (427.81%). Overall, the bias and errors are
significantly lower for OCM2 than in other datasets (MODIS and VIIRS). The slope, coefficient of determination ($R^2$), and root mean square of the regression (RMSE) are listed in Table 2 for the scatter plots of satellite Chl-$a$ data with in situ data (Fig. 7).

The high $R^2$ and slopes close to unity suggest that the validation comparisons are in agreement with the measured dynamic range. The slope values are significantly deviating from the unity in the case of three regression equations, i.e., MODIS (3.73), VIIRS (1.25), and OCM2 (0.40). The highest slope value was noticed for the MODIS algorithm due to the widespread data (0.28 to 18.75 mg m$^{-3}$). The $R^2$ values for the three regression equations are not so strong but show significant variability ($p < 0.001$), out of which the highest value was observed for MODIS ($R^2 = 0.36$), followed by OCM2 ($R^2 = 0.32$) and VIIRS ($R^2 = 0.19$) (Table 2 and Fig. 7). Out of the three sensors studied, the MODIS can explain the higher 36.0% variability of in situ Chl-$a$, followed by OCM2 (32.0%) and VIIRS (19.0%).

The scattered plots symbolize the wide distribution of regression points around the 1:1 line. The expression of RMSE for the log-transformed data of three different sensors concerning in situ values depicted the overall uncertainty and observed highest for MODIS (0.411), VIIRS (0.284), and lower for OCM2 (0.126) (Table 2).

### Discussion

The satellite-derived maps provide a unique temporal and spatial picture of the Chl-$a$ at a global scale (McClain et al., 1998a, 1998b). The OBPG has developed methodologies to validate many historical and operational ocean color missions and the datasets, in such cases a high-quality data set covering a multi-year period is a must and the results can be used for global climate studies (Bailey & Werdell, 2006). The global accuracy in the case of clear and natural waters of modern sensors for the Chlorophyll-$a$ was defined in the range of 5% to 35%, (Hooker et al., 1992). Marina et al. (2006) have reported that the surface

---

**Table 2** Validation statistics of the matchups of Chlorophyll-$a$ (mg m$^{-3}$) performed in the northeastern Arabian Sea

| Parameter | $C_{\text{a,MODIS}}$ vs. $C_{\text{a,in-situ}}$ | $C_{\text{a,OCM2}}$ vs. $C_{\text{a,in-situ}}$ | $C_{\text{a,VIIRS}}$ vs. $C_{\text{a,in-situ}}$ |
|-----------|------------------------------------------|------------------------------------------|------------------------------------------|
| Ratio (± SIQR) | 2.53 (± 1.42) | 0.90 (± 0.27) | 1.75 (± 0.79) |
| MPD (%) | 192.39 | 35.01 | 142.94 |
| Bias | 294.02 | 34.37 | 154.88 |
| RMSE | 427.81 | 49.01 | 212.72 |
| Log_bias | 0.348 | − 0.031 | 0.200 |
| Log_RMSE | 0.411 | 0.126 | 0.284 |
| $a$ | − 1.69 | 0.76 | 0.68 |
| $b$ | 3.73 | 0.40 | 1.25 |
| $R^2$ | 0.36 | 0.32 | 0.19 |
Chlorophyll-\textit{a} concentration (C\textsubscript{a}, mg m\textsuperscript{-3}) in the Southern Ocean was significantly lower than in situ water samples and also revealed the percentage overestimation of in situ data would increase with the decrease in concentrations. However, satellite observations are limited to the ocean surface, and the error was calculated by matchup analysis of concurrent satellite and in situ measurements, which were evaluated to vary around \(\pm 35\%\) in the open ocean (Moore et al., 2009). Antoine et al. (1996) classify the geographic diversity into three trophic regimes based on the Chl-\textit{a} concentration, i.e., oligotrophic (\(C\textsubscript{a} \leq 0.1\) mg m\textsuperscript{-3}), mesotrophic (\(0.1 < C\textsubscript{a} \leq 1.0\) mg m\textsuperscript{-3}) and eutrophic (\(C\textsubscript{a} > 1.0\) mg m\textsuperscript{-3}). The in situ Chl-\textit{a} measurements represented the northeastern Arabian Sea falls under the two spans, i.e., mesotrophic and eutrophic (\(0.1 > C\textsubscript{a} > 1.0\) mg m\textsuperscript{-3}), which explains the region with better productivity. Previous research investigations stated that the northeastern part of the Arabian Sea is one of the highly productive zones and sustains above-average biological productivity (Chauhan et al., 2001; de Sousa et al., 1996; Dwivedi et al., 2006; Shamugam, 2011; Smitha et al., 2022). The present study region is located amid two gulfs, i.e., the Gulf of Kutch and the Gulf of Kambath, which supports better primary productivity because of the vertical mixing of nutrients. The possible reasons for the high productivity due to the coastal upwelling by monsoonal winds (Swallow, 1984) and convective mixing of dense north waters by northeast monsoons (Kumar & Prasad, 1996).

In the study region, the range of Chl-\textit{a} concentration was, i.e., in situ (0.50 to 2.96 mg m\textsuperscript{-3}), MODIS (0.28 to 18.75 mg m\textsuperscript{-3}), VIIRS (0.36 to 11.82 mg m\textsuperscript{-3}) and OCM2 (0.57 to 2.38 mg m\textsuperscript{-3}), which are in similar ranges with the previous studies. The Chl-\textit{a} data were acquired from encompassing, which is ranging from 0.012 to 72.12 mg m\textsuperscript{-3} (Werdell & Bailey, 2005). The OC2v2 algorithm was used in the marginal west sea of the Antarctic Peninsula, for which the SeaWiFS imagers underestimate Chlorophyll-\textit{a} (0.7–43 mg m\textsuperscript{-3}) concentrations by 60\% (Dierssen & Smith, 2000). Along the South Georgia region (54.5°S, 37.0°W) underestimation of 87\% was found in the lower Chl-\textit{a} concentration range (0–1.0 mg m\textsuperscript{-3}), but a 30\% underestimation was recorded for the Chl-\textit{a} concentration value of > 5 mg m\textsuperscript{-3} (Kwok & Comiso, 2002). The Chl-\textit{a} concentration ranges from 0.5 to 7.1 mg m\textsuperscript{-3}, with an average value of 2 mg m\textsuperscript{-3} in the Godavari estuarine waters of India (Latha et al., 2013). The Chl-\textit{a} concentration of satellite data with in situ values, the OCM2, had a similar range of Chl-\textit{a} concentration in the region.

The permissible errors in the log-transformed C\textsubscript{a} data during the algorithm development were about 0.2 or less, and nearly 60\% root mean square (RMS) relative error (O’Reilly et al., 2000). In the global validation of C\textsubscript{a}, the error for most of the ocean basins is about 0.3 (Gregg and Casey 2004), but in the case of the Southern Ocean, the reported errors are significantly larger. According to Raman (2013), the RMSE for the log-transformed data of satellite-derived versus in situ Chl-\textit{a} value was higher for OC2 (23.50\%), OC4 (20.90\%), and OC-OCM (16.10\%). The present results, i.e., log bias and log RMSE percentage, are coherent with the previous studies for different algorithms MODIS (log bias is 0.348 and log RMSE is 41.10\%); VIIRS (log bias is 0.200 and log RMSE is 28.40\%) and OCM2 (log bias is – 0.031 and log RMSE is 12.60\%). The OCM2 sensor performed better for the estimation of Chl-\textit{a} in the region with lower log RMSE and log bias values than the other two sensors, i.e., MODIS and VIIRS. A comparative analysis of Chl-\textit{a} during bloom and non-bloom periods of phytoplankton was studied along Ardley Cove, King George Island, and observed as maximum Chlorophyll-\textit{a} reached 9.87 mg m\textsuperscript{-3} during the bloom period. But, the average Chlorophyll-\textit{a} concentration observed was less than 2 mg m\textsuperscript{-3} from 1991 to 2009. In the case of high latitude and high Chlorophyll-\textit{a} regions, the fluorescence approach ensures a better correlation at 59.46\% (Chen et al., 2017). The satellite measurements still have large errors in estimating phytoplankton biomass (Marrari et al., 2006) and the global Chlorophyll-\textit{a} satellite algorithm typically underestimates Chlorophyll-\textit{a} in the Southern Ocean (Friedrichs et al., 2009; Gregg & Casey, 2004).

In many cases, it remains obvious to underestimate (> 2.0 mg m\textsuperscript{-3} Chlorophyll-\textit{a}) and overestimates (< 2.0 mg m\textsuperscript{-3} Chlorophyll-\textit{a}), but the RMS was surpassed by 60\% for both OC3M and OC3V algorithms. The Chlorophyll-\textit{a} data from MODIS and VIIRS sensors showed a weak correlation with in situ Chl-\textit{a}. However, the VIIRS sensor still performs better than MODIS in the estimation of Chl-\textit{a} with a better regression coefficient. The little differences in band ratio were estimated by using a sensitivity test while declaring the selection of both MODIS and VIIRS (Chen et al., 2016). The RMSE percentage measures the similarity/dissimilarity between the in situ and satellite-derived Chl-\textit{a} value, the OCM2 had lower RMSE (49.01\%) than the other algorithms, i.e., VIIRS (212.72\%) and MODIS (427.81\%). Various empirical algorithms have been developed successfully using regional data sets by considering the local geophysical conditions (Darecki and Stramski 2004; Garcia et al., 2005; Gohin et al., 2002). The current research attempt is to strengthen the in situ database in support of the above objective to develop practical algorithms.

The regression analysis showed significant variability in comparison with in situ data, i.e., MODIS \((R^2 = 0.36\) and slope = 3.73), followed by OCM2 \((R^2 = 0.32\) and slope = 3.73).
validation of remotely sensed products in comparison with in situ values. The accuracy performance of different sensors studied in the region, the OCM2 showed better overall statistical indicators, i.e., MPD, RMSE, Bias, log bias, and log RMSE and regression coefficients than the other sensors (MODIS and VIIRS). The global sensors (MODIS and VIIRS) overestimate in situ Chl-\(a\) concentration (mg m\(^{-3}\)), while the OCM2 algorithm is in tune with the in situ values in the northeastern Arabian Sea. Overall, the satellite sensors (MODIS and VIIRS) systematically overestimated Chl-\(a\) concentration in comparison with in situ data, mainly due to the significant influence of sediments, CDOM, and land discharges. The validation exercise in the region suggests better accuracy indicators for the OCM2 algorithm and is more suitable for the region’s bio-optical characteristics.

**Conclusions**

The Northern Arabian Sea is considered the most productive region in the globe and with unique bio-geochemical and oceanographical coupling processes. Because of the region’s productivity identity and scarcity of in situ data, the current study was initiated to generate in situ Chl-\(a\) datasets and also to evaluate the performance of ocean color satellite sensors in the region. In obvious, the validation process of the satellite datasets is coupled with many limitations, i.e., atmospheric correction due to cloud cover and sun glint, lack of large in situ dataset, the disparity in the in situ data analysis methodologies, coverage of satellites in spatial and temporal scales and frequency of matchup points, etc. Among three satellite sensors (MODIS, VIIRS, and OCM2) evaluated in the region, the OCM2 sensor dataset performed better than the others (the ratio is 0.97 ± 0.27 mg m\(^{-3}\); bias is 34.37%; RMSE is 49.01%; slope 0.40 and \(R^2\) is 0.32 (\(p < 0.001\)). In relation to in situ data, the MODIS (2.53 mg m\(^{-3}\)) and VIIRS (1.75 mg m\(^{-3}\)) sensors indicate the overestimation of Chl-\(a\) concentration in the region, while the OCM2 (0.97 mg m\(^{-3}\)) sensor match with the ground truth data. The error bias in the case of the OCM2 sensor is comparatively lower than the global error accuracy limit (35%) and is suggested to use in the region. However, large in situ datasets are required to calibrate regional-specific algorithms to estimate Chl-\(a\) more accurately in the northeastern Arabian Sea. In consideration of the field-level practical difficulties, the in situ sampling protocols (Standard Operating Procedures, SOPs) need to be formulated as per the international standards (JGOFS) among the diverse organizations engaged in the field of validation experiments.
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12524-022-01598-5.

Acknowledgements Authors are thankful to the Indian Council of Agricultural Research (ICAR), Dr. A. Gopalakrishnan, Director, ICAR-CMFRI, Dr. Mohammed Koya, and Dr. Divu, Scientist-In-Charges, Veraval Regional Centre and Dr. Sathianandan T.V., Head, Fishery Resources Assessment Division, CMFRI, Cochin for the encouragement and support during the study period. The authors would like to thank Dr. Mini Raman, Scientist (SG), ISRO-Space Application Center for the useful technical input pertains to the study region.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Al-Naimi, N., Dionysios, E., Raittso, R. B. H., & Yousria, S. (2017). Evaluation of satellite retrievals of chlorophyll-a in the Arabian Gulf. Remote Sensing, 2017(9), 301. https://doi.org/10.3390/rs9030301

Antoine, D., Andre, J. M., & Morel, A. (1996). Oceanic primary production: 2. Estimation at a global scale from satellite (CZCS) chlorophyll. Global Biogeochemical Cycles, 10, 57–70.

Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sensing of Environment, 102(2006), 12–23.

Baneck, K., & English, D. C. (2000). Geographical differences in seasonality of CZCS-derived phytoplankton pigment in the Arabian Sea for 1978–1986. Deep-Sea Research Part II, 47(7–8), 1623–1677.

Barbini, R., Colal, F., Fantoni, R., Fiorani, L., Okladnikov, I. G., & Palucci, A. (2005). Comparison of SeaWiFS, MODIS-Terra, and MODIS-Aqua in the Southern Ocean. International Journal of Remote Sensing, 26(11), 2471–2478.

Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. Journal of Geophysical Research, 100, 13237–13254.

Chauhan, P., Nagur, C. R., Mohan, M., Nayak, S. R., & Navalgund, R. R. (2001). Surface chlorophyll distribution in Arabian Sea and Bay of Bengal using IRS-P4 ocean color monitor satellite data. Current Science, 80, 127–129.

Chen, Z., Huiping, X., & Andrew, M. F. (2016). Chlorophyll-a estimation around the antarctica peninsula using satellite algorithms: hints from field water leaving reflectance. Sensors, 16, 2075. https://doi.org/10.3390/s16122075

Chen, Z., Tao, Z., Andrew, M. F., & Huiping, X. (2017). Fluorescence-based approach to estimate the Chlorophyll-a concentration of a phytoplankton bloom in aridley cove (Antarctica). Remote Sensing, 9, 210. https://doi.org/10.3390/rs9030210

Clark, D. K. (1981). Phytoplankton pigment algorithms for the Nimbus-7 CZCS. In J. F. R. Gower (Ed.), Oceanography from space (pp. 227–237). Plenum Press.

Claustre, H., Antoine, D., Boehme, L., Boss, E., D’Ortenzio, F., Fantaon D’Andon, O., Guinet, C., Gruber, N., Handegard, N. O., Hood, M., Johnson, K., Lampt, R., LeTraon, P.-Y., Lequere, C., Lewis, M., Perry, M.-J., Platt, T., Roemmich, D., Testor, P., Yoder, J. (2010). Guidelines towards an integrated ocean observation system for ecosystems and biogeochemical cycles. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), Proceedings of OceanObs’09: Sustained Ocean Observations and Information for Society (Vol.2), Venice, Italy, 21–25 September 2009. ESA Publication. https://doi.org/10.5270/OceanObs09

de Sousa, S. N., Dileepkumar, M., Sardessai, S., Sarma, V. V. S. S., & Shirokdar, P. V. (1996). Seasonal variability in oxygen and nutrients in the central and eastern Arabian Sea. Current Science, 71, 847–851.

Dierssen, H. M., & Smith, R. C. (2000). Bio-optical properties and remote sensing ocean color algorithms for Antarctic Peninsula waters. Journal of Geophysical Research, 2000(105), 26301–26312.

Dwivedi, R. M., Raman, M., Parab, S., Matondkar, S. G. P., & Nayak, S. (2006). Influence of northeasterly trade winds on intensity of winter bloom in the northern Arabian Sea. Current Science, 90(10), 1397–1406.

Esaias, W. E., Abbott, M. R., Barton, I., Brown, O. B., Campbell, J. W., & Carder, K. L. (1998). An overview of MODIS capabilities for ocean science observations. IEEE Transactions on Geoscience and Remote Sensing, 36, 1250–1265.

Friedrichs, M. A., Carr, M. E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Buitenhuis, E. T., & Chai, F. (2009). Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. Journal of Marine Systems, 2009(76), 113–133.

Gabric, A. J., Shepard, J. M., Knight, J. M., Jones, G., & Trevena, A. J. (2005). Correlations between the satellite-derived seasonal cycles of phytoplankton biomass and aerosol optical depth in the Southern Ocean: Evidence for the influence of sea ice. Global Biogeochemical Cycles, 19, GB4018.

Garcia, C. A. E., Garcia, V. M. T., & McClain, C. R. (2005). Evaluation of SeaWiFS chlorophyll algorithms in the Southwestern Atlantic and Southern Oceans. Remote Sensing of Environment, 95, 125–137.

Gohin, F., Druon, J. N., & Lampert, L. (2002). A five-channell chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. International Journal of Remote Sensing, 23, 1639–1661.

Grieg, W. W., & Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll-a data set. Remote Sensing of Environment, 2004(93), 463–479.

Grieg, W. W., & Conkright, M. (2001). Global seasonal climatologies of ocean chlorophyll: Blending in situ and satellite data for the Coastal Zone Color Scanner era. Journal of Geophysical Research: Oceans, 106, 2499–2515.

Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., & Strickland, J. D. H. (1965). Fluorometric determination of chlorophyll. ICES Journal of Marine Science, 30(3–15), 1965. https://doi.org/10.1093/icesjm/30.1.3

Hooker, S. B., Esaias, W. E., Feldman, G. C., Greg, W. W., & McClain, C. R. (1992). An overview of SeaWiFS and ocean color: NASA Tech. Memo., vol. 104566.

Hooker, S. B., Van Heuken, L., Thomas, C. S., Claustre, H., Ras, J., Schultler, L., Clements, L., Van der Linde, L., Eker-Develi, E., Berthon, J.-F., Barlow, R., Sessions, H., Ismail, H., & Perl, J. (2009). The third SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE–3). NASA technical Memorandum 2009-215849. NASA Goddard Space Flight Center.

IOC. (2000). Remote sensing of ocean colour in coastal, and other optically-complex waters. In S. Sathyendranath (Ed.), Reports of
Werdell, P. J., & Bailey, S. W. (2005). An improved bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment, 98*, 122–140.

Zhang, C., Hu, C., Shang, S., Müller-Karger, F. E., Li, Y., & Dai, M. (2006). Bridging between SeaWiFS and MODIS for continuity of chlorophyll a concentration assessments off southeastern China. *Remote Sensing of Environment, 102*, 250–263.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.