

# Retrieval of Chlorophyll-a and vegetation indices Using Sentinel-2 MSI Imagery in El-Burullus Lake, Egypt

Walaa Assar<sup>1, □</sup>



**Abstract** Chlorophyll-a (Chl-a) is considered an important indicator for assessing many environmental issues. Monitoring Chl-a accurately through remote sensing technology is an efficient approach to observing water quality and providing early warnings of water eutrophication which overcomes the limitations in field campaigns as well as cost constraints. The Sentinel-2 Multi-Spectral Imager (S2-MSI) was used to analyse El-Burullus Lake, which is considered the second-largest natural lake in Egypt. The Case-2 Regional/Coast Colour (C2RCC) atmospheric correction processor was applied to retrieve different optically active water constituents such as Chl-a, the total suspended matter and coloured dissolved organic matter, as well as vegetation indicators. The spatial distribution marked the eutrophicated areas about 87.2 % of El-Burullus Lake area which result from the inflow from a large number of polluted agricultural drains. The results of the validation between the captured S2-MSI Chl-a and in situ measurements confirmed a clear underestimation in the level of the retrieved C2RCC parameters as on average of 65.1 % for the Chl-a values. The low  $R^2$  values (below 0.1481) mean the bio-optical model better fits the variations in the in situ Chl-a content, as the F-test showed a significant level for all the parameters. All the results, which included low values for the mean bias error (MBE), root mean square error and normalized root mean square error; below 15.13, 28.05, and 1.005 values, respectively, demonstrated that the empirical model exhibits a good performance with the bio-

-optical model, and the positive MBE results indicated the underestimation in the in situ Chl-a data. Accordingly, the S2-MSI imageries proved their applicability for the efficient mapping, water quality assessment and sustainable protection of the lake ecosystem.

**Keywords:** Chlorophyll-a; El-Burullus Lake; Sentinel-2; Statistical Analysis; Vegetation Indicators.

## 1 Introduction

Lakes and wetlands have ecological importance due to their role in biodiversity conservation, as they provide valuable socioeconomic services for fisheries and birds. Multiple human activities such as land reclamation, untreated wastewater discharge, and aquaculture have an adverse influence on their conservation potential globally [1] [2] [3]. Chlorophyll-a (Chl-a), the photosynthetic pigment in phytoplankton, is considered the main indicator for estimating many environmental issues such as water quality, algae biomass, and eutrophication [4]. The algae abundance and its correlated Chl-a concentrations fluctuate spatially and temporally in lakes, which makes monitoring in situ periodic Chl-a levels difficult due to the limited financial resources and spatial constraints, in addition to the significant time required [5].

In this context, remote sensing (RS) technology has an efficient ability for monitoring environmental changes in large-scale areas, as well as water pollution and vegetation activity, that overcomes the spatiotemporal limitations [6]. Thus, the retrieval of Chl-a concentrations and the accurate estimation of its controlling parameters, in terms

Received: 18 June 2023/ Accepted: 31 October 2023

□ Corresponding Author, Walaa Assar,

[walaa.assar@f-eng.tanta.edu.eg](mailto:walaa.assar@f-eng.tanta.edu.eg)

1.Irrigation and Hydraulics Engineering Department, Faculty of Engineering, Tanta University, 31511 El-Gharbia, Egypt

[Orcid ID: 0000-0002-4545-1636](https://orcid.org/0000-0002-4545-1636)

of the total suspended matter (TSM) and the coloured dissolved organic matter (CDOM), by RS has significant spatiotemporal advantages for effective water quality assessment [7]. The recent generation of high-resolution multispectral sensors, the Sentinel-2 Multi-Spectral Imager (S2-MSI), can significantly enhance the spatial resolution by 10 m, 20 m and 60 m in 13 spectral bands (every five days), distributed along the visible/near-infrared (VNIR) and short-wave infrared (SWIR) spectral bands for more accuracy [8] [9]. Therefore, Sentinel-2 is considered the best solution as it demonstrates good spectral resolution in capturing Chl-a levels, which is suitable for Chl-a estimation [10].

Numerous algorithms have been proposed for estimating optical constituents in water (Case-2 water), including empirical (statistical regression or end member selection), analytical (involving the mechanism of the radiative transfer equation in water), and semi-analytical (a combination between empirical and analytical) approaches [11]. Other ones, such as the bio-optical model, were initially developed to detect the optical properties of the water including phytoplankton and its breakdown substances based on the spectral measurements obtained from satellite images [12]. In Case-2 water (inland and coastal water bodies), the retrieval of Chl-a is more complex and less accurate because of the interaction between Chl-a, TSM, and CDOM [13]. Additionally, for Chl-a retrieval, it is necessary to enhance the atmospheric correction, as it is influenced by cloud cover [14].

The Case-2 Regional/Coast Colour (C2RCC) atmospheric correction processor has proved its capability in relation to Sentinel-2 data using a set of neural networks which determines the water-leaving radiance from the top-of-atmosphere (TOA) radiances, as well as the retrieval of the inherent optical properties (IOPs) of the water body [15]. Thus, remotely sensed water constituents (Chl-a, TSM, and CDOM) for analysing and monitoring are necessary to obtain an accurate estimation of the spatiotemporal water quality distribution, which results in a better understanding and sustainable protection of aquatic ecosystems.

With the development of multispectral and hyperspectral RS technology, vegetation indices have been proposed to estimate the quantitative and qualitative assessment of vegetation cover, growth, water content and pigment composition, based on the electro-magnetic spectrum from canopies [16]. Three different vegetation indices were applied to derive the Chl-a content as an important bio-indicator from RS data, including the green normalized difference vegetation index (GNDVI),

maximum chlorophyll index (MCI), and pigment-specific simple ratio (PSSRa).

Therefore, the main objective of this research was to assess the water quality status based on Sentinel-2 images through the retrieval of Chl-a, TSM, and CDOM concentrations.

For this purpose, empirical and bio-optical models were developed and validated with in situ Chl-a measurements, during which a statistical analysis was performed to examine the quantitative relationships between the water constituents and vegetation indices.

## 2 Material and Methods

### 2.1 Study Area

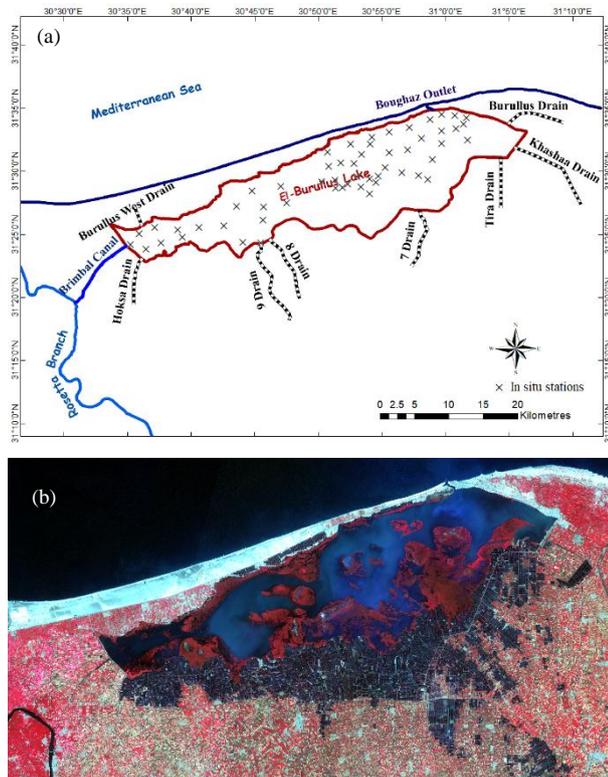
These El-Burullus Lake, a shallow brackish lake considered the second-largest natural one in Egypt, displays typical lagoon characteristics, which extend along the Mediterranean Nile Delta of Egypt. Therefore, periodic monitoring of the ecosystem's water quality is essential for protecting it from deterioration. El-Burullus Lake lies between the two main Nile River branches (the Rosetta and Damietta), extending between longitudes 30° 30' E and 31° 10' E and latitudes 31° 21' N and 31° 35' N (Fig. 1a). It covers an area of approximately 455 km<sup>2</sup>, with a length of approximately 54 m parallel to the Mediterranean coast, a width varying between 3 and 12 km, and water depths ranging from 0.4 to 2.5 m.

The lake water body is discharged through eight agricultural drains, as well as being connected to fresh water from the Nile through the Brimbil Canal on the western side [17]. El-Burullus Lake receives approximately 3904 million m<sup>3</sup>/year of agricultural, industrial, domestic, and fish farm wastewater [18].

Numerous studies over the last four decades have indicated that the lake is suffering from a dramatic deterioration and aquatic surface area reduction, due to as the increasing irregular urbanization and intensive anthropogenic activities, as well as more recently the impact of climate change, which have enhanced the eutrophication and vegetation growth [19] [20] [21] [22]. Moreover, RS technique was applied for mapping water quality parameters combined with field measurements to investigate its applicability of assessment spatial and temporal changes along El-Burullus Lake [23] [24] [25]. For instance, Mohsen in 2021 applied field observations from (August 2010 to August 2013), in 2022, Hossen utilized data of 2015, and in situ measurements on September 2020 were carried out by Masoud in 2021.

Therefore, the recent one was selected to assess the water quality status based on Sentinel-2 images.

Figure 1b illustrates the Sentinel-2 MSI False-colour Infrared (RBG) of El-Burullus Lake on 18 September 2020, showing denser vegetation (B8) in red and water (B3) in blue.

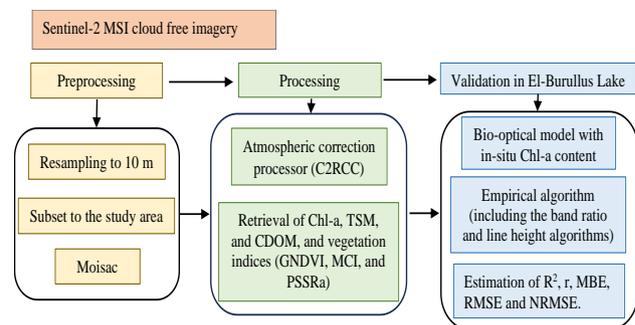


**Fig. 1** Map of El-Burullus Lake (a) location of the study area including connections to agricultural drains and the Nile River, and In-situ stations, (b) Sentinel-2 MSI False-colour Infrared image on 18 September 2020

## 2.2 Sentinel-2 Data and Image Processing

The Sentinel-2 mission provides global coverage of the Earth's surface every five days with a spatial resolution of 10 to 60 m, and this short repetition cycle results in high-resolution optical imagery that can be used for mapping the surface water changes and capturing Chl-a variations. For the present study, freely available Sentinel-2 images were downloaded from the US Geological Survey (USGS) website Earth Explorer (<https://earthexplorer.usgs.gov/>). Sentinel-2 (Sentinel-2A, S2A/MSI) cloud-free imagery covering El-Burullus Lake was utilized for the analysis on 18 September 2020, as the

only available cloud-free imagery that is nearly simultaneous with the in situ measurements which carried out during the period 22 to 25 September 2020. With the [Sentinel](#) Application Platform (SNAP) software, which is offered without charge and has been developed to facilitate its utilization, the images were resampled to a 10 m (B2 band) and subset to the mapped lake, displaying and processing the remotely sensed data [26]. Then, the atmospheric correction processor (C2RCC) was applied to produce the corrected normalized water-leaving radiances at bands 1 to 6 ( $\rho_{\text{w}}^{\text{corr}}(\lambda)$ ), which allows for the retrieval of different optically active water constituents such as Chl-a, TSM, and CDOM [27]. A flow-chart diagram of the research process and methodology is shown in Fig. 2.



**Fig. 2** Flow-chart diagram of the research methodology

## 2.3 Vegetation Indices

Vegetation and pigment indicators, including the green normalized difference vegetation index (GNDVI), maximum chlorophyll index (MCI), and pigment-specific simple ratio (PSSRa), were applied as remote sensing indices to identify the different concentration rates of chlorophyll.

### 2.3.1 Green Normalized Difference Vegetation Index (GNDVI)

The GNDVI was initially proposed by Gitelson et al. [28], who verified that GNDVI was resistant to atmospheric effects and more sensitive to chlorophyll

content in a wide range of chlorophyll variations.

It has been widely applied as the most common index for determining nitrogen uptake in plant canopy and pigment concentration studies [29].

The GNDVI represents the vegetation's chlorophyll concentration, which is derived from the difference between near-infrared radiation (NIR) (terrestrial

vegetation) and the green spectral (GREEN) reflectance of the electromagnetic spectrum [30] [31], with a range of - 1 to 1, as shown in Equation 1:

$$GNDVI = \frac{(NIR - GREEN)}{(NIR + GREEN)} \quad (1)$$

### 3.2 Maximum Chlorophyll Index (MCI)

The MCI algorithm exploits the height of a measurement in a certain spectral band (B5, 705 nm) over a specific baseline, which passes through two other spectral bands (B4, 665 nm and B6, 740 nm) in Sentinel-2 MSI data [32]. The MCI algorithm implemented in the S2 MCI processor for floating vegetation and inland water bodies utilizes the TOA conditions as follows:

$$MCI = L_2 - k \times \left( L_1 + (L_3 - L_1) \times \frac{(\lambda_2 - \lambda_1)}{(\lambda_3 - \lambda_1)} \right) \quad (2)$$

Where L 1 and 3 indicate the spectral baseline wavelengths (B4, B6), L 2 the peak spectral band (B5), and  $\lambda_i$  centred at wavelength, while the k factor corrects the effect of thin clouds [33].

#### 2.3.3 Pigment-Specific Simple Ratio (PSSRa)

The PSSRa algorithm was developed by Blackburn [34] to investigate the potential of a range of spectral approaches for quantifying pigments at the scale of the whole plant canopy.

The PSSRa represents the strongest and most linear relationships with the canopy concentration per unit of area of chlorophyll-a (Chl-a), chlorophyll-b (Chl-b) and the carotenoids (Cars) by applying a narrow-band pigment. It is derived from the simple ratio between the red edge (RED-EDGE) (central wavelength) B7 (783 nm) and red spectral radiation (RED) bandwidth B4 (665 nm) in Sentinel-2 data [35].

### 2.4 Validation Assessment

Regression analyses with the coefficient of determination ( $R^2$ ) and the Pearson correlation coefficient (r) were applied to evaluate the models' performance, using

Minitab software. Moreover, the mean bias error (MBE), root mean square error (RMSE), and normalized root mean square error (NRMSE), as basic statistical indicators, were employed to quantify the forecasting differences and the dispersion between the estimated and in situ data. In Equations 3-5, low values indicate better model performance.

$$MBE = \frac{1}{N} \left[ \sum_{i=1}^n \frac{(E_i - O_i)^2}{N} \right] \quad (3)$$

$$RMSE = \sqrt{\left( \frac{\sum_{i=1}^n (E_i - O_i)^2}{N} \right)} \quad (4)$$

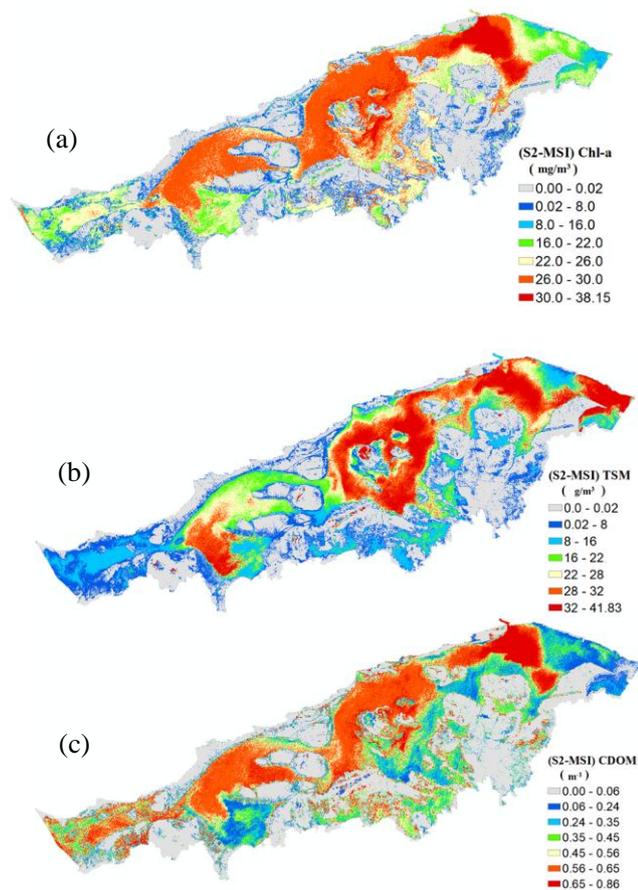
$$NRMSE = \frac{RMSE}{\left( \frac{\sum_{i=1}^n O_i}{N} \right)} \quad (5)$$

where, N = number of observations, E = estimated value, and O = observed value.

## 3 Results and Discussion

### 3.1 Water Constituent Retrievals

The distributions of the retrieved water parameters of Chl-a, TSM, and CDOM along El-Burullus Lake were extracted as optically active constituents by the atmospheric correction processor (C2RCC) in the SNAP program for processing the S2-MSI images on 18 September 2020. The mosaic of processed images was extracted to the lake using geographic information system (GIS) software. Figure 3 shows high levels of Chl-a, TSM, and CDOM towards El-Boughaz and the centre of the lake parallel to the coastline. In addition to the abovementioned locations, more local highs were observed in front of Drain 7 for Chl-a content (Fig. 3a), in the eastern region next to the El-Burullus and El-Gharbia drains for TSM values (Fig. 3b), and additional increases in CDOM along the western area (Fig. 3c). The averages of the retrieved water constituents were 21.91 mg/m<sup>3</sup>, 19.13 g/m<sup>3</sup>, and 0.463 m<sup>-1</sup>, for Chl-a, TSM, and CDOM, respectively. The content ranged from 0.0037 to 38.15 mg/m<sup>3</sup> of Chl-a, 0.0095 to 41.83 g/m<sup>3</sup> of TSM, and 0.0002 to 0.8576 m<sup>-1</sup> of CDOM. These bloom areas of Chl-a, TSM, and CDOM are intensified by the discharge of large numbers of polluted agricultural drains into the lake.

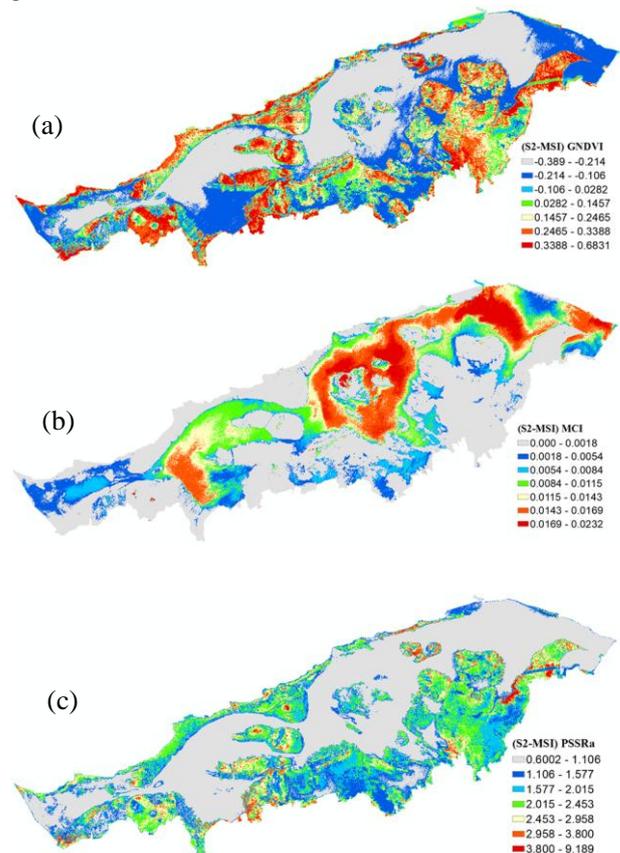


**Fig. 3** Sentinel-2 MSI retrieval images of water constituents (a) Chl-a, (b) TSM, and (c) CDOM along El-Burullus Lake on 18 September 2020

### 3.2 Vegetation and Pigment Indicators

Spectral vegetation and pigment indicators including GNDVI, MCI, and PSSRa were calculated for biomass valuations. Figure 4a shows the spatial distribution of GNDVI with a good resolution in Sentinel-2 MSI images along El-Burullus Lake. The GNDVI ranges from -1 to 1, where the positive values represent the vegetation canopy areas, and the negative ones indicate the absence of vegetation. It is resistant to atmospheric corrections, since it is five times more sensitive to MSI Chl-a content, with a good correlation ( $r = -0.882$ ), and a linear regression ( $R^2 = 0.777$ ). The high levels of GNDVI were observed towards the northern and the southern borders of the lake, especially next to the Tira, 7, 8, 9, and Hoksa drains. Figure 4b illustrates high values of MCI near El-Boughaz, the El-Burullus Drain and the centre of the lake parallel to the coastline, which corresponded to the distribution of Chl-a in Fig. 3a ( $r = 0.699$ ), with a linear regression ( $R^2 =$

0.646). The PSSRa was applied not only to estimate the vegetation biomass but also for monitoring its health across the lake through the interpretation of the active Chl-a content variations (Fig. 4c). The PSSRa distribution has a good correlation with the Chl-a content and GNDVI ( $r = -0.843$  and  $0.956$ ) as well as a linear regression ( $R^2 = 0.711$  and  $0.914$ ). The negative MSI Chl-a correlation with GNDVI and PSSRa could be due to the difference in the behaviour of Chl-a and the vegetation indices.



**Fig. 4** Sentinel-2 MSI images showing vegetation indices (a) GNDVI, (b) MCI, and (c) PSSRa along El-Burullus Lake on 18 September 2020

### 3.3 Validation of the Bio-Optical Model

The bio-optical model was validated by deriving the Chl-a content from a dataset of 55 in situ measurements distributed along El-Burullus Lake during 20 to 25 September 2020, as shown in Fig. 1a. The physicochemical and biological analyses for the lake were performed in a recent study by Masoud et al. [25], who reported that the water status of El-Burullus Lake is hypertrophic, with wide variations in Chl-a content,

ranging from 53.76 to 129.03 mg/m<sup>3</sup>. Chl-a content was analysed spectrophotometrically according to Mackinney [36], at the laboratories of the National Institute of Oceanography and Fisheries (NIOF), Alexandria (Egypt).

A statistical summary of the in situ Chl-a concentrations is illustrated in Table 1. The results indicated the bloom areas of Chl-a, which intensified in front of the eastern drains, as well as Drain 7 and Brimbil Canal in the west. A regression analysis of the in situ Chl-a data was performed against the derived Sentinel-2 MSI Chl-a, CDOM, TSM, and the vegetation indices GNDVI, MCI, and PSSRa. Due to the large difference between the observed Chl-a values (average = 83.43 mg/m<sup>3</sup>) and the derived Sentinel-2 MSI Chl-a concentrations (average = 21.91 mg/m<sup>3</sup>), weak to fair correlation coefficients are presented in Table 2.

Moreover, the determination coefficient (R<sup>2</sup>) scored low values near to zero, which clarified that the bio-optical model is not reliable for predicting the in situ Chl-a outcomes (Fig. 5). However, high values of R<sup>2</sup> do not always relate to the strength of the regression analysis, because the quality of R<sup>2</sup> depends on various factors, such as the small number of in situ data samples which were utilized and the fact that the samples were not taken at the same time as the images. Chl-a values are very sensitive to variations in time and space.

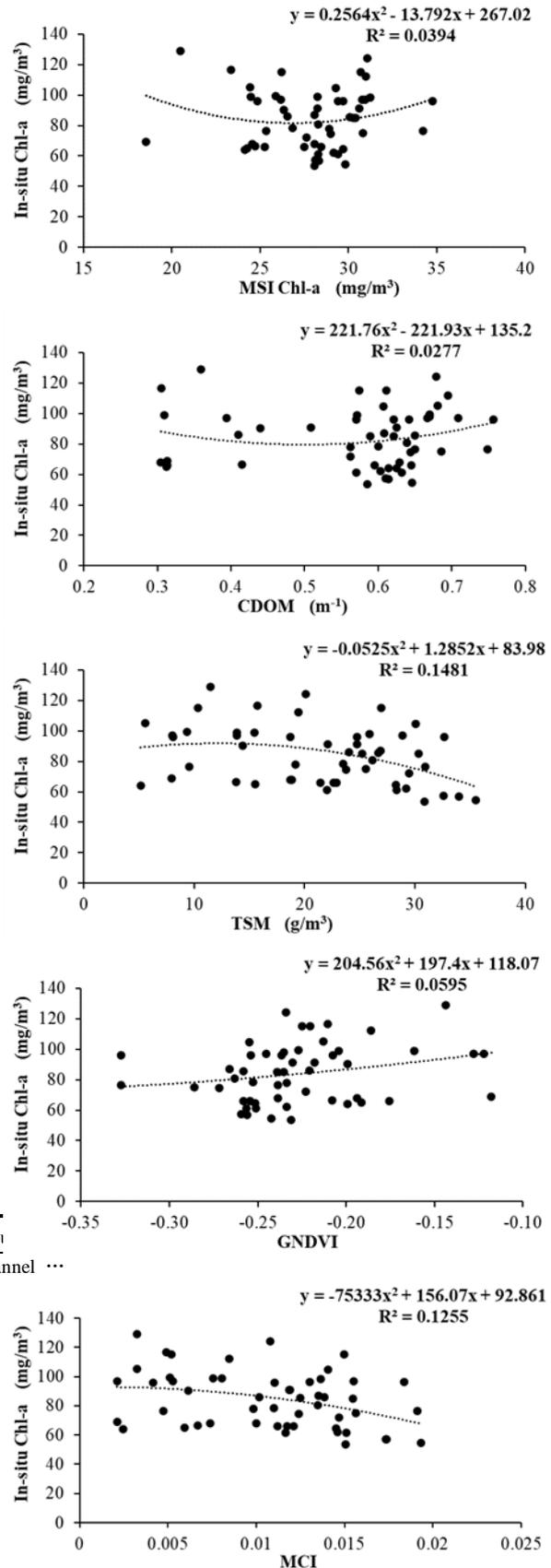
Therefore, the F-test (hypothesis test) was applied to determine whether the relationship between the model and the in situ data was statistically significant. The results of the overall significance F-test demonstrated a significant level (P-value < 0.05) for all parameters. Thus, the bio-optical model better fits the variations of in situ Chl-a content.

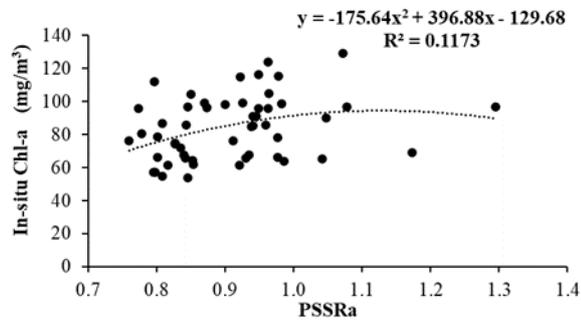
**Table 1** Statistics summary of the in situ Chl-a (mg/m<sup>3</sup>) content along El-Burullus Lake

Min	Max	Mean	Std. error	Stand. dev	Groi
53.76	129.03	83.43	2.57	19.09	Channel ...
25 Percentil	75 Percentil	Skewness	Kurtosis		
66.05	97.03	0.39	-0.61		

**Table 2** Channels arrangement

	MSI Chl-a	CDOM	TSM	GNDVI	MCI	PSSRa
In Situ Chl-a	0.013	0.014	-0.342	0.242	-0.343	0.299





**Fig. 5** Regression analysis of the in situ data Chl-a content against the derived Sentinel-2 MSI Chl-a, CDOM, TSM, GNDVI, MCI, and PSSRa

An empirical algorithm has been proposed for estimating and validating the empirical relationship between the in situ Chl-a values and the water-leaving radiances in sensor bands, including the band ratio and line height algorithms. The band ratio algorithm was applied to retrieve the Chl-a concentration, utilizing the NIR-red ratio, which is typically employed in turbid and CDOM enriched water bodies [37].

The NIR-red ratio uses the near-infrared band between the 700 and 720 wavelengths, as well as the red band which is located in the range of maximum absorption for Chl-a between the 660 and 690 wavelengths, as displayed in Equation 6, developed by [38].

$$Chl_a \sim \frac{R(705)}{R(665)} \rightarrow \frac{B5}{B4} \quad (6)$$

The line height algorithm, as the maximum chlorophyll index (MCI), was derived from the detection of surface blooming vegetation in wetlands, coastal areas and oceans [39], as in Equation 7:

$$MCI = R(705) - R(665) - 0.53 \times (R(740) - R(665)) \\ \rightarrow B5 - B4 - 0.53 \times (B6 - B4) \quad (7)$$

Where the value 0.53 is an index representing the ratio of wavelengths for the applied bands.

Figure 6 illustrates the regression model of the in situ Chl-a values against the empirical algorithms, the band ratio (R705/R665), and the calculated MCI. The  $R^2$  scored low values, 0.032 and 0.126, and a negative Pearson's correlation coefficient ( $r = 0.286$  and  $0.428$ ) for the band ratio and calculated MCI, respectively. However, a high  $R^2$  was noted between the bio-optical model and the

empirical model of the Chl-a estimation model, as shown in Fig. 7. The  $R^2$  values of the MSI Chl-a content against the band ratio and calculated MCI were 0.2164 and 0.4776, with high positive correlations ( $r = 0.462$  and  $0.690$ ), respectively.

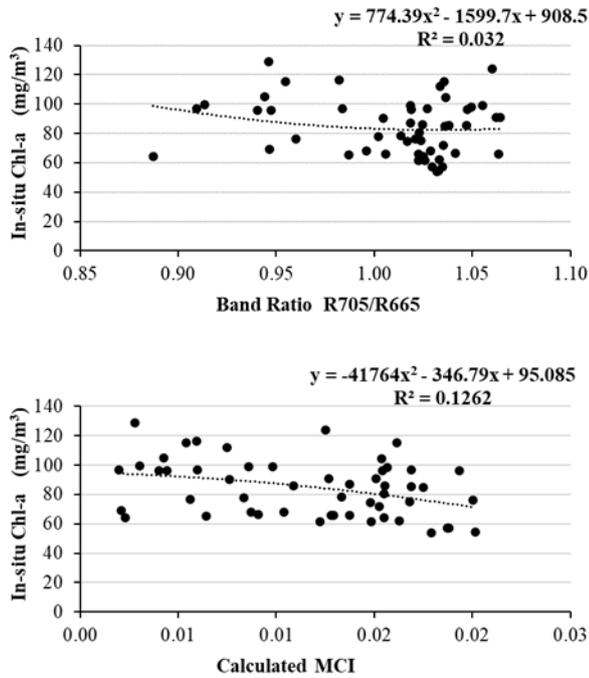
To validate the performance of the in situ Chl-a values and the bio-optical model with the empirical model of the Chl-a estimation, the MBE, RMSE, and NRMSE were calculated. In Table 3, the MBE values were 130.59 and 14.08, with RMSE values of 82.41 and 27.06, and NRMSE ones of 0.993 and 0.970, for the in situ Chl-a and the bio-optical model with the band ratio, respectively. However, for the in situ Chl-a and the bio-optical model with the line height algorithm, the values of MBE were 133.75 and 15.13, the RMSE ones were 83.40 and 28.05, and the NRMSE ones were 1.004 and 1.005, respectively.

All the outcomes with low values of MBE, RMSE, and NRMSE demonstrate that the empirical model has a good performance with the bio-optical model, and the positive MBE results indicate the underestimation in the in situ Chl-a data.

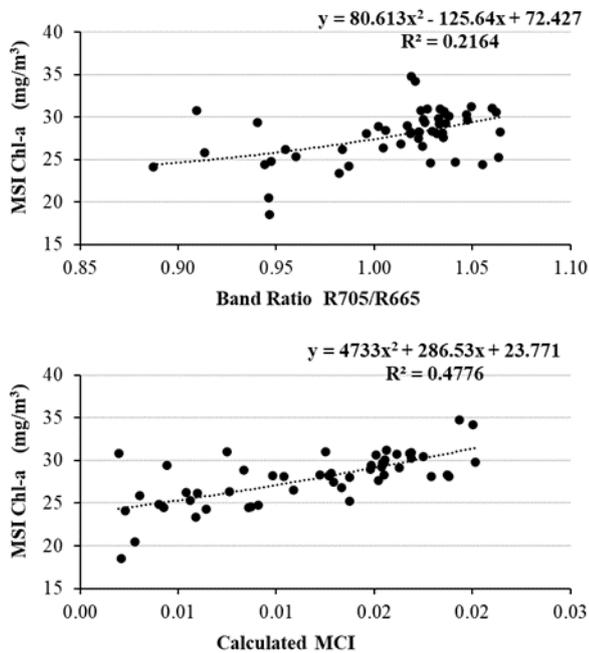
Based on the outcomes in this study, the S2-MSI imageries proved their applicability for mapping and monitoring the bloom vegetation areas through retrieved water constituents. The spatial distribution of Chl-a, TSM, CDOM and the vegetation indices marked the eutrophicated areas along El-Burullus Lake which result from the inflow from large numbers of polluted agricultural drains into the lake. These bloom areas are consistent with the findings of previous works which indicated the areas of deterioration due to the intensive inflow from polluted drains along the lake [40] [7].

Thus, S2-MSI data can be applied for efficient mapping and water quality monitoring for the assessment, better understanding and sustainable protection of the lake ecosystem, overcoming the limitations in field campaigns as well as cost constraints.

However, the C2RCC algorithm failed to present accurate water constituents, underestimating the Chl-a content in a shallow and hypertrophic lake (< 2m depth). This could be attributed to turbidity and high values of TSM and CDOM, where CDOM absorbs at an increasing rate in the blue bands and masks the Chl-a absorption in the blue regions. The C2RCC presents the active optical properties in the visible and near-infrared wavelengths, indicated the abundance of algae, which consistent with the bloom areas of previous studies. These results confirm those of [15] [41] and [42], who reported the validity of the C2RCC algorithm for spatiotemporal Chl-a changes in inland waters.



**Fig. 6** Regression analysis of the in situ data Chl-a content against the empirical algorithms, the band ratio (R705/R665), and the calculated MCI



**Fig. 7** Regression analysis of the derived MSI Chl-a concentration from the optical model against the empirical algorithms, the band ratio (R705/R665), and the calculated MCI

**Table 3** Statistical indicators of the in situ Chl-a and the derived MSI Chl-a (mg/m<sup>3</sup>) with the empirical algorithms, the band ratio (R705/R665), and the calculated MCI.

In Situ Chl-a	R <sup>2</sup>	r	MBE	RMSE	NRMSE
Band Ratio	0.032	-0.286	130.59	82.41	0.993
Calculated MCI	0.126	-0.428	133.75	83.40	1.004
MSI Chl-a					
Band Ratio	0.2164	0.462	14.08	27.06	0.970
Calculated MCI	0.4776	0.690	15.13	28.05	1.005

#### 4 Conclusion

This study used high-resolution optical imagery from Sentinel-2 (S2A/MSI) for mapping and monitoring Chl-a variations along El-Burullus Lake, which is considered the second-largest natural lake in Egypt, on 18 September 2020. C2RCC was applied as an atmospheric correction processor to retrieve different optically active water constituents such as Chl-a, TSM, and CDOM utilizing the SNAP program. The mosaic of processed images was extracted to the lake with GIS software. The results of the distribution of Chl-a, TSM, and CDOM illustrated high levels towards El-Boughaz and the centre of the lake parallel to the coastline. Moreover, the bloom areas are intensified by the discharge of large amounts of polluted agricultural drains into the lake.

The average values were 21.91 mg/m<sup>3</sup>, 19.13 g/m<sup>3</sup>, and 0.463 m<sup>-1</sup>, for Chl-a, TSM, and CDOM, respectively, and ranged from 0.0037 to 38.15 mg/m<sup>3</sup> of Chl-a, 0.0095 to 41.83 g/m<sup>3</sup> of TSM, and 0.0002 to 0.8576 m<sup>-1</sup> of CDOM. Additionally, the vegetation indicators (GNDVI, MCI and PSSRa) were utilized as remote sensing indices to identify different concentration rates of chlorophyll. The GNDVI is resistant to atmospheric corrections, as it is five times more sensitive to MSI Chl-a content, with a good correlation ( $r = -0.882$ ), and a linear regression ( $R^2 = 0.777$ ). The high levels of GNDVI were noted towards the northern and the southern borders of the lake, especially next to the Tira, 7, 8, 9, and Hoksa drains. High values of MCI were observed near El-Boughaz, the El-Burullus Drain and the centre of the lake parallel to the coastline, which correspond to the distribution of Chl-a ( $r = 0.699$ ), with a linear regression ( $R^2 = 0.646$ ).

Moreover, the PSSRa distribution has a good correlation with the Chl-a content and GNDVI ( $r = -0.843$  and  $0.956$ ) and a linear regression ( $R^2 = 0.711$  and  $0.914$ ). The negative MSI Chl-a correlation with GNDVI and PSSRa could be due to the difference between the dependent variables in the behaviour of Chl-a and the vegetation indices.

For the bio-optical model validation, the Chl-a content was measured from 55 in situ data measurements recorded along El-Burullus Lake during 20 to 25 September 2020. The results of the validation between the retrieved Sentinel-2 MSI Chl-a and in situ concentrations confirmed a clear underestimation in the level of the retrieved C2RCC parameters as on average of 65.1 % for the Chl-a values. Despite the low values of  $R^2$ , the bio-optical model better fits the variations of in situ Chl-a content, and the F-test showed a significant level ( $P\text{-value} < 0.05$ ) for all parameters. Secondly, the proposed empirical algorithms, including the band ratio and line height ones, were employed to calculate the Chl-a concentrations as well as the MCI. Three performance statistics (MBE, RMSE, and NRMSE), were applied to the in situ Chl-a values and the bio-optical model with the empirical model for Chl-a estimation. All the results, with low values for MBE, RMSE, and NRMSE, demonstrated that the empirical model exhibits a good performance with the bio-optical model, and the positive MBE results indicate the underestimation in the in situ Chl-a data.

Accordingly, the S2-MSI imageries proved the instrument's applicability for mapping and monitoring the bloom vegetation areas through the retrieved water constituents. The spatial distribution of Chl-a, TSM and CDOM as well as the vegetation indices marked the eutrophicated areas along El-Burullus Lake which result from the large amounts of inflow from polluted agricultural drains into the lake. Thus, S2-MSI data can be applied for efficient mapping and water quality monitoring for the assessment, better understanding and sustainable protection of the lake ecosystem, overcoming the limitations in field campaigns as well as cost constraints. However, the C2RCC algorithm failed to present accurate water constituents, underestimating the Chl-a content in a shallow and hypertrophic lake ( $< 2\text{m}$  depth).

## References

- [1] K. H. Shaltout and Y. M. Al-Sodany, "Vegetation analysis of Burullus Wetland: a RAMSAR site in Egypt," *Wetl. Ecol. Manag.*, vol. 16, no. 5, pp. 421–439, 2008.
- [2] Y. Zhang et al., "Global loss of aquatic vegetation in lakes," *Earth-Science Rev.*, vol. 173, pp. 259–265, 2017.
- [3] L. Szabó, B. Deák, T. Bíró, G. J. Dyke, and S. Szabó, "NDVI as a proxy for estimating sedimentation and vegetation spread in artificial lakes—Monitoring of spatial and temporal changes by using satellite images overarching three decades," *Remote Sens.*, vol. 12, no. 9, p. 1468, 2020.
- [4] X. Shi, L. Gu, T. Jiang, X. Zheng, W. Dong, and Z. Tao, "Retrieval of Chlorophyll-a Concentrations Using Sentinel-2 MSI Imagery in Lake Chagan Based on Assessments with Machine Learning Models," *Remote Sens.*, vol. 14, no. 19, p. 4924, 2022.
- [5] T. R. B. T. Aranha, J.-M. Martinez, E. P. Souza, M. U. G. Barros, and E. S. P. R. Martins, "Remote Analysis of the Chlorophyll-a Concentration Using Sentinel-2 MSI Images in a Semiarid Environment in Northeastern Brazil," *Water*, vol. 14, no. 3, p. 451, 2022.
- [6] F. Pei, Y. Zhou, and Y. Xia, "Application of normalized difference vegetation index (NDVI) for the detection of extreme precipitation change," *Forests*, vol. 12, no. 5, p. 594, 2021.
- [7] A. A. Masoud, "On the Retrieval of the Water Quality Parameters from Sentinel-3/2 and Landsat-8 OLI in the Nile Delta's Coastal and Inland Waters," *Water*, vol. 14, no. 4, p. 593, 2022.
- [8] S. Liang, Z. Gong, Y. Wang, J. Zhao, and W. Zhao, "Accurate monitoring of submerged aquatic vegetation in a macrophytic lake using time-series sentinel-2 images," *Remote Sens.*, vol. 14, no. 3, p. 640, 2022.
- [9] L. Abad, D. Hölbling, R. Spiekermann, G. Prasicek, Z. Dabiri, and A.-L. Argentin, "Detecting landslide-dammed lakes on Sentinel-2 imagery and monitoring their spatio-temporal evolution following the Kaikōura earthquake in New Zealand," *Sci. Total Environ.*, p. 153335, 2022.
- [10] H. Liu et al., "Eutrophication monitoring of lakes in Wuhan based on Sentinel-2 data," *GIScience Remote Sens.*, vol. 58, no. 5, pp. 776–798, 2021.
- [11] I. Moutzouris-Sidiris and K. Topouzelis, "Assessment of Chlorophyll-a concentration from Sentinel-3 satellite images at the Mediterranean Sea using CMEMS open source in situ data," *Open Geosci.*, vol. 13, no. 1, pp. 85–97, 2021.
- [12] M. H. Gholizadeh, A. M. Melesse, and L. Reddi, "A comprehensive review on water quality parameters estimation using remote sensing techniques," *Sensors*, vol. 16, no. 8, p. 1298, 2016.
- [13] R. Doerffer and H. Schiller, "The MERIS Case 2 water algorithm," *Int. J. Remote Sens.*, vol. 28, no. 3–4, pp. 517–535, 2007.
- [14] X. Hang, Y. Li, X. Li, M. Xu, and L. Sun, "Estimation of Chlorophyll-a Concentration in Lake Taihu from Gaofen-1 Wide-Field-of-View Data through a Machine Learning Trained Algorithm," *J. Meteorol. Res.*, vol. 36, no. 1, pp. 208–226, 2022.
- [15] M. Pereira-Sandoval et al., "Evaluation of atmospheric correction algorithms over Spanish inland waters for sentinel-2 multi spectral imagery data," *Remote Sens.*, vol. 11, no. 12, p. 1469, 2019.
- [16] E. G. Tafesse, T. D. Warkentin, S. Shirtliffe, S. Noble, and R. Bueckert, "Leaf Pigments, Surface Wax and Spectral Vegetation Indices for Heat Stress Resistance in Pea," *Agronomy*, vol. 12, no. 3, p. 739, 2022.
- [17] W. Assar, M. Elshemy, and B. A. Zeidan, "Water Quality Modeling for Lake Burullus, Egypt, Part I: Model Calibration.(Dept. C (irrigation))," *MEJ. Mansoura Eng. J.*, vol. 40, no. 2, pp. 54–61, 2020.
- [18] M. S. Moussa and M. K. Mostafa, "Rapid Assessment Method for Evaluation of the Weighted Contribution of Anthropogenic Pollution: A Case Study of Lake Burullus, Egypt," *Water*, vol. 13, no. 23, p. 3337, 2021.

- [19] M. A. El-Alfy, D. H. Darwish, and Y. A. El-Amier, "Land use Land cover of the Burullus Lake shoreline (Egypt) and health risk assessment of metal-contaminated sediments," *Hum. Ecol. Risk Assess. An Int. J.*, vol. 27, no. 4, pp. 898–920, 2021.
- [20] A. Shalby, M. Elshemy, and B. A. Zeidan, "Modeling of climate change impacts on Lake Burullus, coastal lagoon (Egypt)," *Int. J. Sediment Res.*, vol. 36, no. 6, pp. 756–769, 2021.
- [21] M. E. A. El-Metwally, D. H. Darwish, and M. A. Dar, "Spatial distribution and contamination assessment of heavy metals in surface sediments of Lake Burullus, Egypt," *Arab. J. Geosci.*, vol. 14, no. 1, pp. 1–13, 2021.
- [22] A. Hany, F. Akl, M. Hagra, and A. Balah, "Assessment of recent rehabilitation projects' impact on water quality improvement in Lake Burullus, Egypt," *Ain Shams Eng. J.*, vol. 13, no. 1, p. 101492, 2022.
- [23] A. Mohsen, M. Elshemy, and B. Zeidan, "Water quality monitoring of Lake Burullus (Egypt) using Landsat satellite imageries," *Environ. Sci. Pollut. Res.*, vol. 28, pp. 15687–15700, 2021.
- [24] H. Hossen, W. E. Mahmood, A. Negm, and T. Nakamura, "Assessing water quality parameters in burullus lake using sentinel-2 satellite images," *Water Resour.*, vol. 49, no. 2, pp. 321–331, 2022.
- [25] A. A. Masoud, M. M. El-Horiny, H. M. Khairy, and M. M. El-Sheekh, "Phytoplankton dynamics and renewable energy potential induced by the environmental conditions of Lake Burullus, Egypt," *Environ. Sci. Pollut. Res.*, vol. 28, pp. 66043–66071, 2021.
- [26] B. Ghansah, T. Foster, T. P. Higginbottom, R. Adhikari, and S. J. Zwart, "Monitoring spatial-temporal variations of surface areas of small reservoirs in Ghana's Upper East Region using Sentinel-2 satellite imagery and machine learning," *Phys. Chem. Earth, Parts A/B/C*, vol. 125, p. 103082, 2022.
- [27] S. Li et al., "Mapping the trophic state index of eastern lakes in China using an empirical model and Sentinel-2 imagery data," *J. Hydrol.*, vol. 608, p. 127613, 2022.
- [28] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 289–298, 1996.
- [29] P. Shanmugapriya, K. R. Latha, S. Pazhanivelan, R. Kumaraperumal, G. Karthikeyan, and N. S. Sudarmanian, "Spatial prediction of leaf chlorophyll content in cotton crop using drone-derived spectral indices," *Curr. Sci.*, vol. 123, no. 12, p. 1473, 2022.
- [30] L. Rodríguez-López et al., "Spectral analysis using LANDSAT images to monitor the chlorophyll-a concentration in Lake Laja in Chile," *Ecol. Inform.*, vol. 60, p. 101183, 2020.
- [31] M. Martello, J. P. Molin, M. C. F. Wei, R. Canal Filho, and J. V. M. Nicoletti, "Coffee-Yield Estimation Using High-Resolution Time-Series Satellite Images and Machine Learning," *AgriEngineering*, vol. 4, no. 4, pp. 888–902, 2022.
- [32] M. Elhag, I. Gitas, A. Othman, J. Bahrawi, and P. Gikas, "Assessment of water quality parameters using temporal remote sensing spectral reflectance in arid environments, Saudi Arabia," *Water*, vol. 11, no. 3, p. 556, 2019.
- [33] J. F. R. Gower, R. Doerffer, and G. A. Borstad, "Interpretation of the 685nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS," *Int. J. Remote Sens.*, vol. 20, no. 9, pp. 1771–1786, 1999.
- [34] G. A. Blackburn, "Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches," *Remote Sens. Environ.*, vol. 66, no. 3, pp. 273–285, 1998.
- [35] N. M. Reddy, I. A. Peerzada, M. Moonis, and O. Singh, "Application of biophysical, soil, and vegetation indices to better understand forest dynamics and develop strategies for forest conservation," in *Forest Dynamics and Conservation: Science, Innovations and Policies*, Springer, 2022, pp. 399–415.
- [36] G. Mackinney, "Absorption of light by chlorophyll solutions," *J. Biol. Chem.*, vol. 140, no. 2, pp. 315–322, 1941.
- [37] M. A. Warren et al., "Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters," *Remote Sens. Environ.*, vol. 225, pp. 267–289, 2019.
- [38] K. G. Ruddick, H. J. Gons, M. Rijkeboer, and G. Tilstone, "Optical remote sensing of chlorophyll a in case 2 waters by use of an adaptive two-band algorithm with optimal error properties," *Appl. Opt.*, vol. 40, no. 21, pp. 3575–3585, 2001.
- [39] C. Zeng and C. Binding, "The effect of mineral sediments on satellite chlorophyll-a retrievals from line-height algorithms using red and near-infrared bands," *Remote Sens.*, vol. 11, no. 19, p. 2306, 2019.
- [40] M. Elshemy, B. A. Zeidan, and W. Assar, "Water Quality Mitigation Scenarios for Burullus Coastal Lake, Egypt," in *Estuaries and Coastal Zones in Times of Global Change*, Springer, 2020, pp. 89–110.
- [41] M. Niroumand-Jadidi, F. Bovolo, L. Bruzzone, and P. Gege, "Inter-comparison of methods for chlorophyll-a retrieval: Sentinel-2 time-series analysis in Italian lakes," *Remote Sens.*, vol. 13, no. 12, p. 2381, 2021.
- [42] I. Ogashawara et al., "The use of Sentinel-2 for chlorophyll-a spatial dynamics assessment: A comparative study on different lakes in northern Germany," *Remote Sens.*, vol. 13, no. 8, p. 1542, 2021.