

## Research Article

# Modelling Reservoir Chlorophyll-*a*, TSS, and Turbidity Using Sentinel-2A MSI and Landsat-8 OLI Satellite Sensors with Empirical Multivariate Regression

Yashon O. Ouma <sup>1,2</sup>, Kimutai Noor,<sup>1</sup> and Kipkemoi Herbert<sup>1</sup>

<sup>1</sup>Department of Civil and Structural Engineering, Moi University, 30100 Eldoret, Kenya

<sup>2</sup>Department of Civil Engineering, University of Botswana, Private Bag 0061, Gaborone, Botswana

Correspondence should be addressed to Yashon O. Ouma; [yashon\\_o@hotmail.com](mailto:yashon_o@hotmail.com)

Received 4 June 2020; Revised 30 July 2020; Accepted 31 August 2020; Published 19 September 2020

Academic Editor: Lei Zhang

Copyright © 2020 Yashon O. Ouma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sentinel-2A/MSI (S2A) and Landsat-8/OLI (L8) data products present a new frontier for the assessment and retrieval of optically active water quality parameters including chlorophyll-*a* (Chl-*a*), suspended particulate matter (TSS), and turbidity in reservoirs. However, because of their differences in spatial and spectral samplings, it is critical to evaluate how well the sensors are suited for the seamless generation of the water quality parameters (WQPs). This study presents results from the retrieval of the WQP in a reservoir from L8 and S2A optical sensors, after atmospheric correction and standardization through band adjustment. An empirical multivariate regression model (EMRM) algorithmic approach is proposed for the estimation of the water quality parameters in correlation with *in situ* laboratory measurements. From the results, both sensors estimated Chl-*a* concentrations with  $R^2$  of greater than 70% from the visible green band for L8 and a combination of green and SWIR-1 bands for S2A. While the NMSE% was nearly the same for both sensors in Chl-*a* estimation, the RMSE was  $<10 \mu\text{g/L}$  and  $>10 \mu\text{g/L}$  for L8 and S2A estimations of Chl-*a*, respectively. For TSS retrieval, L8 outperformed S2A by 31% in accuracy with  $R^2 > 0.9$  from L8's red, blue, and green bands, as compared to  $0.47 \leq R^2 \leq 0.61$  from S2A's red and NIR bands. The RMSE were the same as for Chl-*a*, and the NMSE% were both in the same range. Both sensors retrieved turbidity with high and nearly equal accuracy of  $R^2 > 70\%$  from the visible and NIR bands, with equal RMSE at  $<10\%$  NTU and NMAE% from S2A being higher by more than 30% as compared to L8's NMAE% at 15%. The study concluded that the higher performance accuracy of L8 is attributed to its higher SNR and spectral bandwidth placement as compared to S2A bands. Comparatively, S2A overestimated Chl-*a* and turbidity but performed equally well compared to OLI in the estimation of TSS. The results show that while absolute accuracy of retrieval of the WQPs still requires improvements, the developed algorithms are broadly able to discern the biooptical water quality in reservoirs.

## 1. Introduction

Despite the fact that there is an increase in the availability of safe drinking water globally, it is still estimated that approximately an eighth of the global population does not have access to potable water [1]. According to [2], nearly half of the world's population will face water shortages by 2030 as the water demand in certain countries will likely to exceed

supply by over 50%. Most of the water sources are within the river and reservoir catchments which are continuously affected by anthropogenic pressure and climate change.

Within a catchment system, the anthropogenic water uses for domestic purposes, agricultural and industrial production, mining, power generation, and forestry practices all result in the deterioration of water quality and water supply. This in effect impacts the aquatic ecosystem and compromises the

access to safe water supply for human consumption. Water quality and water supply are thus linked, although they are often not measured simultaneously. The results of water quality monitoring are important in determining the spatio-temporal trends in surface water and groundwater variabilities. For reservoirs and dams, as the key source of water for cities and rural communities, the *a priori* understanding of the quality of water before treatment is critical not only in understanding the environmental health within the catchment area but also in the minimization of the cost of treatment of the water supply.

To monitor the water quality in natural and artificial reservoirs, the conventional water quality assessment through sampling and laboratory measurement is often employed. The sampling point-based laboratory methods are costly, labor-intensive, and time-consuming and are not able to adequately assess the entire water body [3]. To overcome the limitations in *in situ* water quality monitoring methods, there is a need for regular near real-time [4–7], inexpensive, automated, and noninvasive approaches, with adequate spatial-temporal coverage.

Several studies have investigated the use of different satellite sensors for the assessment of water quality. Particularly, the Landsat sensors have been widely used in the estimation of water quality parameters such as total suspended matter, chlorophyll-*a*, turbidity, Secchi disk depth, total phosphorus, dissolved oxygen, chemical oxygen demand (COD), and biochemical oxygen demand (BOD) as reviewed in [8, 9]. For specific case studies and for the retrieval of water quality parameters (WQPs), the previous studies have developed different correlational algorithms that are based on empirical models, semianalytical models, and matrix inversion models. For semianalytical models, both the biooptical and empirical data are required to describe the relationships between the constituents of a water body and the equivalent surface reflectance that defines the upwelling radiance above the water surface and in the water surface. Based on the same modelling scheme as the semianalytical models, the matrix inversion models require *a priori* information on the water constituents, including the absorption coefficients or absorption slopes [10]. The lack of specific parameters makes the matrix inversion methods more complex and difficult to calibrate. Because of these drawbacks, empirical algorithms are often used for the retrieval and estimation of water quality parameters [3, 11–14].

Though popular, in the use of empirical algorithms, large water quality sample sizes are required and the models are sensitive to local environmental conditions and are therefore not automatically replicable to other case studies or regions. In addition to being case study or region based, the empirical models have been developed more for open sea waters as compared to inland water bodies, due in part to the fact that remote sensing measurements of freshwater resources are far more complex, in terms of surface water spectral reflectivity [15–17].

With the potential of higher temporal resolution of about 2–3 days derived from the synergistic constellation of Landsat-8 and Sentinel-2 [18], it is now possible to synchronize the products from the two satellites' higher temporal

monitoring of aquatic systems [19]. Such frequent revisits are essential for the capture of the dynamics of reservoir water bodies, in terms of surface water quality assessments, considering the effects of seasonal variabilities and atmospheric attenuations. In using Landsat-8/OLI and Sentinel-2A/MSI data for water resource monitoring, recent studies have presented different approaches and results for different case study areas. [20] used Sentinel-2A to study the influence of variations in the concentration of total suspended solids (TSS) and chlorophyll-*a* (Chl-*a*) on the physiological response of oysters and highlighted the use of Sentinel-2A near-infrared (NIR) bands to quantify the total suspended solids. [11] developed empirical-based methods for the retrieval of Chl-*a* in a hypereutrophic reservoir in Brazil and concluded that the Sentinel-2A near-infrared (NIR) bands were significant in Chl-*a* retrieval. Similar studies by [13] presented empirical algorithms using Sentinel-2A data for water quality assessment [21] or bottom mapping [22]. Nonetheless, no attempts have been made to compare and demonstrate the suitability of Sentinel-2A/MSI and Landsat-8/OLI imagery for the retrieval of Chl-*a*, TSS, and turbidity water quality parameters within inland water reservoirs.

In this study, the effectiveness of Sentinel-2A/MSI (S2A) and Landsat-8/OLI (L8) satellite sensors is demonstrated for the estimation of Chl-*a*, turbidity, and TSS water quality parameters in a large reservoir (case 2 water body). Arguably, the medium-spatial resolution satellite sensors, Operational Land Imager (OLI) on board Landsat-8 and Multispectral Imager (MSI) on board Sentinel-2, will be capable of promoting more precise mapping of biooptically active water quality parameters in recent times [11, 19]. However, because of their differences in the spectral and spatial samplings, it is critical to evaluate how well the datasets are suited for the seamless retrieval of water quality parameters. The accuracy of the biooptically active parameters as mapped from different remote sensors is largely dependent on the biooptical equation developed for its retrieval. As observed in the literature review, different case 2 water bodies respond differently to the spectral wavelengths.

Further, for case 2 waters, the concentrations of the water constituents and the corresponding water color are considered to be nonlinear. This implies that for effective measurements in these highly reflective waters, remote sensors with high dynamic spectral ranges and high signal-to-noise ratio (SNR) are needed [17]. It is therefore necessary to develop optimal algorithm(s) for the accurate estimation of biooptical water quality parameters in regional case 2 waters. The current study has two objectives: (1) to identify the most suitable spectral bands (position and width) from the Sentinel-2A/MSI and Landsat-8/OLI sensors for accurate retrieval and estimations of the concentrations of Chl-*a*, TSS, and turbidity and (2) develop, test, and validate empirical multivariate regression model (EMRM) algorithms for the estimation of the water quality parameters in case 2 waters, in correlation with *in situ* laboratory measurements.

To determine the distribution and variability of water quality parameters, ordinary Kriging is used for the spatial mapping and comparison of the distributions of the WQP in the case study reservoir. The rest of the paper is organized

as follows: in Section 2, the study area and the details of the data provided are described. The research methods are outlined in Section 3, and Section 4 presents the study results and discussions which elaborate on the comparison and analyses of the regression modelling and estimation of the WQPs from the compared satellite sensors. The study conclusions and insights are presented in Section 5.

## 2. Study Area and Data

**2.1. Study Site Characterization.** The case study reservoir is the Chebara Dam which is located between longitudes  $35^{\circ}29'45.6''\text{E}$  and  $35^{\circ}30'7.2''\text{E}$  and latitudes  $0^{\circ}52'55.2''\text{S}$  and  $0^{\circ}53'56.4''\text{S}$  and was constructed in the 1990s to supply water to Eldoret Town, which is situated 31 km away. The 2.1 km long dam has a capacity of 6.24 million cubic meters and is served by River Moiben and other small rivers and streams within the Chebara basin (Figure 1). Most of the rivers and streams flowing into the dam are permanent and flow throughout the year, with least flows experienced during the dry seasons. The climate in the Chebara Dam catchment is warm and temperate, and the area receives significant rainfall even in the driest month, with an annual precipitation average of approximately 1172 mm. Sparsely settled, agriculture is the main economic activity within the basin, with the rest of the land cover being forest, grass, and shrubs (Figures 1(a) and 1(b)). The dam also serves as a source of water to the resident population for domestic use and agricultural irrigation, especially during the dry season or periods.

The sampling stations were selected as shown in Figure 1, with the location and concentration of the sampling stations being based on the depth variations of the dam. Eighteen (18) spatially distributed sampling stations were selected: ten (10) for the calibration of the EMRM algorithm, five (5) for validation, and three (3) for off-season algorithm validation.

### 2.2. Optical Satellite Sensors

**2.2.1. Sentinel-2/MSI Sensor.** The Sentinel-2 mission comprises twin polar-orbiting satellites, Sentinel-2A and Sentinel-2B, which were, respectively, launched on 23 June 2015 and 7 March 2017. The sensors are in the same orbit and phased at  $180^{\circ}$  to each other. The orbit is sun-synchronous at 786 km altitude with an orbit inclination of  $98.62^{\circ}$ , and the mean local solar time at the descending node is at 10:30 AM. Both the orbit inclination and timing ensure minimum cloud cover and optimal sun illumination. The Sentinel-2 satellites are also aligned with similar satellites, including Landsat and SPOT-5.

Each Sentinel-2 satellite carries a multispectral instrument (MSI), with a swath width of approximately 290 km and spatial high resolutions of 10 m, 20 m, and 60 m, with a revisit time of 10 days with one satellite and 5 days with two satellites at the equator. The Sentinel MSI has 13 spectral bands (430 nm to 2320 nm) in the visible, red-edge, near-infrared, and shortwave infrared regions of the spectrum. With high-temporal resolution and 12-bit radiometric resolution, like the Landsat-8 sensor, the MSI sensor provides

high radiometric dynamics for the observed areas of both the very dark, e.g., water surfaces, and the very bright areas, e.g., ice and snow. It is further observed that the locationing and bandwidth of Landsat OLI are slightly wider than those of the MSI sensor spectral bands, as also indicated in Table 1 for the bandwidth and their spectral locations.

For comparison with the Sentinel-2A MSI, only the similar spectral bands in Landsat OLI are analyzed as presented in Table 1.

**2.2.2. Landsat-8/OLI Sensor.** The Landsat-8 mission carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). OLI collects image data for nine short-wave bands with spatial sampling or resolution of 30 meters (bands 1-7) and the panchromatic band with pixel size of 15 meters (band 8). TIRS collects data for two long-wave thermal bands at 100 meters (band 9) every 16 days which are resampled to 30 meters to match OLI multispectral bands. The OLI sensor is compatible with the earlier Landsat sensors and presents improved measurement capabilities. Compared to the Landsat-7/ETM+, the Landsat-8/OLI has improved radiometric resolution with reduced image noise and spectral heterogeneity. This is observed to be particularly significant in precise water surface extraction and water quality retrieval [23].

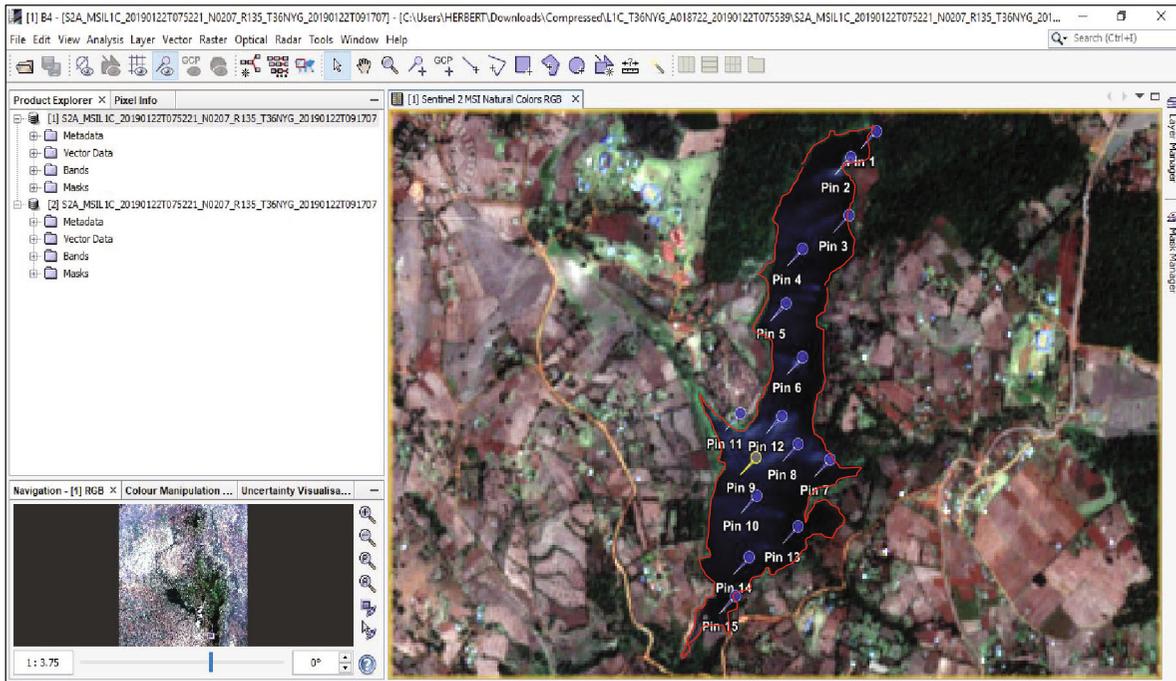
This study utilized the Landsat OLI data acquired on 22 January 2019 and the Sentinel-2A satellite image acquired on 25 January 2019. The fieldwork was carried out during the dry season and coincided with the sensor acquisitions from 23 to 24 January 2019. For the Sentinel-2A, the level 1C is corrected for atmospheric errors using the Sentinel's SNAP algorithm, resulting in the level 2A product of Sentinel-2 from the Sen2Cor processor, which includes scene classification and atmospheric correction.

## 3. Methods

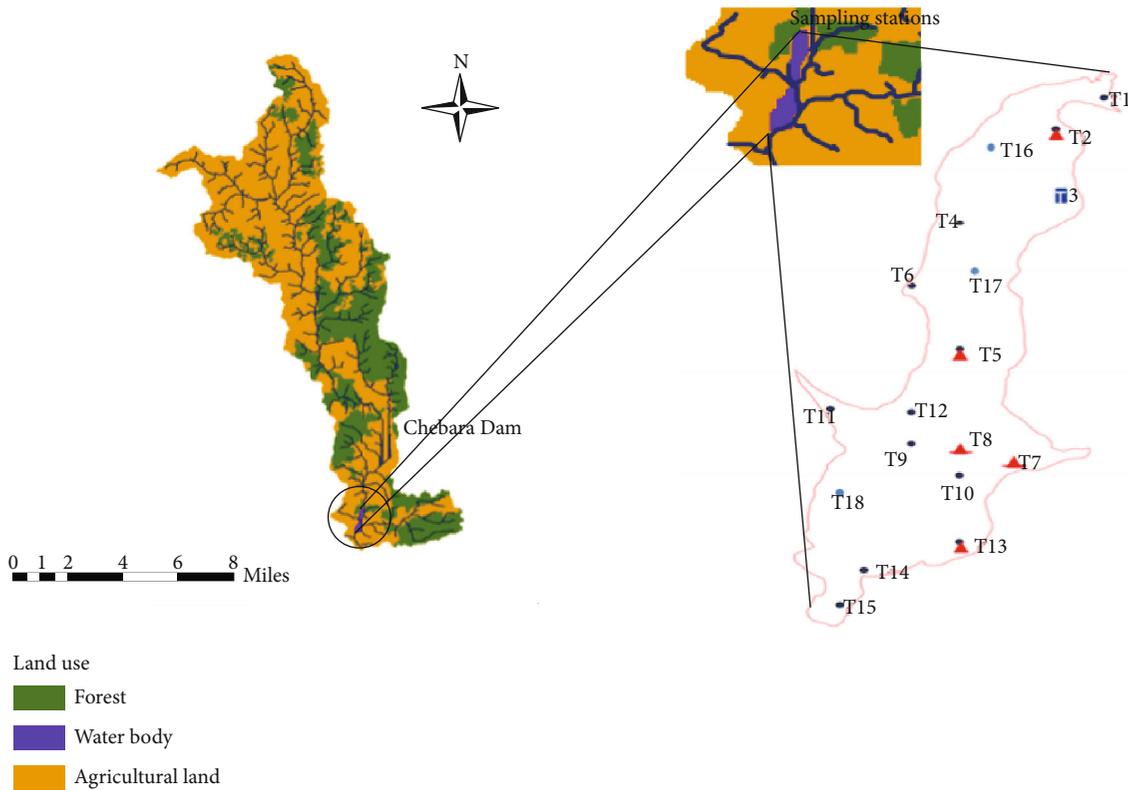
The level 1 data of the Landsat-8 OLI sensor consists of scaled digital number values that are quantized and calibrated. The equivalent level 1 data from Sentinel-2 MSI comprises the top of atmosphere (TOA) reflectance. In this study, the retrieval of the water quality parameters from the satellite sensors over the study region involves the following four steps: (i) deriving the absolute TOA reflectance from scaled DN values for Landsat OLI and scaled TOA reflectance for Sentinel-2 MSI, respectively, for the homologue bands in Table 1; (ii) conversion of TOA reflectance to surface reflectance, which is the actual reflectance originating from the water surface; (iii) conversion of the surface reflectance to the corresponding remote sensing reflectance ( $R_{rs}$ ) at these bands; and (iv) the retrieval of the water quality parameters from  $R_{rs}$  utilizing the proposed empirical multivariate regression model (EMRM).

### 3.1. Preprocessing of Sentinel-2A and Landsat OLI Datasets

**3.1.1. TOA Reflectance Derivation and Signal-to-Noise Ratio Analysis.** To derive and compare the TOA from the optical sensors, for Landsat-8 level 1, the digital numbers from each band ( $B_{L8}$ ) are converted to TOA reflectance using the



(a)



(b)

FIGURE 1: (a) Sentinel-2 near-true color composite of the study dam and surrounding areas and (b) land use/land cover in the Chebara basin and the sampling points ( $T_i$ ).

Landsat-8 metadata scaling factors and divided by the cosine of the solar zenith angle (see processing steps in [3]). In addition, Landsat per-pixel solar and viewing angles were

calculated using the Landsat-8 Angles Creation Tool provided by USGS (<https://landsat.usgs.gov/solar-illumination-and-sensor-viewing-anglecoefficient-file>) and the angle

TABLE 1: Spectral and spatial samplings of the visible, NIR, and SWIR bands in Landsat OLI and Sentinel-2A.

| Band # and spectral range | Landsat-8 OLI            |                |                        | Band # and spectral range | Sentinel-2A MSI          |                |                        |
|---------------------------|--------------------------|----------------|------------------------|---------------------------|--------------------------|----------------|------------------------|
|                           | Spectral resolution (nm) | Bandwidth (nm) | Spatial resolution (m) |                           | Spectral resolution (nm) | Bandwidth (nm) | Spatial resolution (m) |
| B1-coastal aerosol        | 435–451                  | 16             | 30                     | B1-coastal aerosol        | 433–453                  | 21             | 60                     |
| B2-blue                   | 452–512                  | 60             | 30                     | B2-blue                   | 458–523                  | 66             | 10                     |
| B3-green                  | 533–590                  | 57             | 30                     | B3-green                  | 543–578                  | 36             | 10                     |
| B4-red                    | 636–673                  | 37             | 30                     | B4-red                    | 650–680                  | 31             | 10                     |
| B5-NIR                    | 851–879                  | 28             | 30                     | B8-NIR                    | 785–900                  | 106            | 10                     |
|                           |                          |                |                        | B9-NIR narrow             | 855–875                  | 21             | 20                     |
| B6-SWIR-1                 | 1566–1651                | 85             | 30                     | B11-SWIR-1                | 1566–1655                | 91             | 20                     |
| B7-SWIR-2                 | 2107–2294                | 187            | 30                     | B12-SWIR-2                | 2100–2280                | 175            | 20                     |

coefficient file available with each Collection 1 L1TP file. For the Sentinel-2 MSI, the TOA reflectance and solar view geometry are stored in the metadata for every 5 km.

**3.1.2. Surface Reflectance from Sentinel MSI and Landsat OLI.** For the TOA reflectance derivation, the datasets from the sensors were converted to the equivalent surface reflectance using the same algorithm in order to minimize any biases in the derivation of the water quality parameters due to derivation of the reflectance. The LaSRC (Land Surface Reflectance Code) algorithm developed for Landsat and based on the 6S radiative transfer code was used and adopted for Sentinel-2A for TOA-to-surface reflectance conversion [24–26]. As validated by [27], the LaSRC has higher accuracy than the original Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm when applied to Landsat-8 OLI data [24, 26].

From the multisensor bandwidth comparisons in Table 1, it is observed that MSI and OLI have variable spatial and spectral sampling resolutions which may result in variabilities in their spectral sensitivities and responses to the same water surface. To determine the significance of the spectral responses, the uniformly corrected sensor bands are compared for the signal-to-noise ratio (SNR) and top of atmosphere (TOA) reflectance over a spatially uniform and clear water body area.

**3.1.3. Lambertian BRDF-Adjusted Reflectance.** Because most land surfaces do not have a Lambertian Bidirectional Reflectance Distribution Function (LBRDF), the reflectance may vary simply due to factors such as the variations in the solar radiation and viewing angle geometry. Because of BRDF effects, the Landsat and Sentinel-2A swath across the red and near-infrared band reflectances can vary by 0.02–0.06 and 0.06–0.08, respectively [28]. These differences and variations can result in significant noise effects as they are comparable and greater than the sensor calibration errors [29]. For the LBRDF reflectance adjustment, the semiphysical approach developed for Landsat application and demonstrated to also work for Sentinel-2A by [28] was used for the adjustment of the surface reflectance for each comparable band to a 0° nadir view and the observed solar zenith angle [23].

**3.1.4. Multisensor Crossband Adjustment.** For compatibility and comparability of the Landsat OLI and Sentinel-2A bands, the observed differences in the homologue bands (Table 1) are crossadjusted as discussed in [30, 31]. The need for band adjustment is to minimize the effects of the bandwidth differences as depicted in Table 1. From Table 2, the six bands (blue, green, red, NIR, SWIR-1, and SWIR-2) were adjusted using the crosssensor transformation coefficients (Table 2) as derived from the results by [31]. Approximately the same crosssensor adjustment coefficients were obtained [32] in the harmonization of Landsat OLI and Sentinel-2A MSI sensor data bands.

**3.2. Empirical Regression Modelling for Retrieval of Water Quality Parameters.** As presented in the introductory and literature review sections, this study proposes the use of EMRM, which follows also from an earlier approach as presented in [3]. Using EMRM, simulations are carried out to establish the multivariate correlations between the sensor band reflectances and the measured *in situ* water quality variables. Table 3 presents the remote sensing reflectance band combination(s) considered for the EMRM analysis of the Sentinel-2A and Landsat OLI data in the estimation of the water quality parameters. In Table 3,  $i$ ,  $j$ , and  $k$  refer to the sensor bands from the Landsat OLI and Sentinel-2A datasets ( $B_{L8}$  and  $B_{S2A}$ ), before conversion to remote sensing reflectance  $R_{rs}(\lambda)$  and multiband adjustment using the crosssensor band transformations.

The multivariate regression model for estimating the water quality parameters in the reservoir is developed by determining the quantitative relationships between the measured *in situ* water quality parameter and the remote sensing reflectance from the satellite spectral data. The empirical models used in the regression of the *in situ* measurements to the sensor band combinations were comprised of the following model equations [3], where  $R_{rs}(\lambda)$  is the corresponding remote sensing reflectance for  $B_{L8}$  and  $B_{S2A}$  bands and  $a$ ,  $b$ , and  $c$  are the regression model constants.

(a) *Linear.*  $a * R_{rs}(\lambda) + b$ .

(b) *Polynomial.*  $a * R_{rs}(\lambda)^2 + b * R_{rs}(\lambda) + c$ .

TABLE 2: Crosssensor band transformation coefficients for Landsat-8 [31].

| Bands  | Landsat-8 OLI sensor |        | Multisensor adjustment equation  |
|--------|----------------------|--------|--|
|        | Intercept            | Slope  |  |
| Blue   | -0.0107              | 1.0946 | $B_{S2A} = \text{slope} * B_{L8} + \text{intercept}$ ,<br>where $B_{S2A}$ is the Sentinel-2A band and $B_{L8}$ is the Landsat OLI band |
| Green  | +0.0026              | 1.0043 |  |
| Red    | -0.0015              | 1.0524 |  |
| NIR    | +0.0033              | 0.8954 |  |
| SWIR-1 | +0.0065              | 1.0049 |  |
| SWIR-2 | +0.0046              | 1.0002 |  |

TABLE 3: Sentinel-2A MSI and Landsat-8 OLI band combination(s) for water quality parameter retrieval using the proposed EMRM.

| Band combination(s)     | Sentinel-2A bands                          | Landsat OLI bands                       |
|-------------------------|--|---|
| Single bands            | $B_{S2A1} = B_{S2Ai}$                      | $B_{L81} = B_{L8i}$                     |
| Linear band combination | $B_{S2A2} = B_{S2Ai} + B_{S2Aj}$           | $B_{L82} = B_{L8i} + B_{L8j}$           |
| Band ratios             | $B_{2A3} = B_{S2Ai}/B_{S2Aj}$              | $B_{L83} = B_{L8i}/B_{L8j}$             |
| Mixed band combinations | $B_{S2A} = (B_{S2Ai}/B_{S2Aj}) + B_{S2Ak}$ | $B_{L84} = (B_{L8i}/B_{L8j}) + B_{L8k}$ |

$B_{L8}$  and  $B_{S2A}$  are, respectively, the Landsat OLI and Sentinel-2A bands, and  $B_{L81-4}$  and  $B_{S2A1-4}$  are the four band combination models 1-4.

(c) *Logarithmic*.  $a * \log_{10} R_{rs}(\lambda) + b$ .

(d) *Power*.  $a * R_{rs}^b(\lambda)$ .

(e) *Exponential*.  $a * e^{b * R_{rs}(\lambda)}$ .

In the development of the EMRM, ten (10) of the sampling point data were used in the regression modelling in model calibration and five (5) remaining data points (T2, T5, T7, T8, and T13) were used in the validation of the model. Surface water samples from T1 to T15 were sampled in January 2019 during the dry season and those from T16 to T18 were sampled in May after the onset of the raining season. To determine the best-fit model, the correlation of determination  $R^2$  between the predicted model and laboratory-measured water quality parameter is ranked, and the model fit is tested at the entry significance level of  $p = 0.05$  and the removal significance level of 0.10. The  $t$ -test at the confidence level of 95% is used to accept the determined regressive model. The sampling point-based retrieved WQPs are also spatially mapped to determine the spatial distribution of the concentrations of the WQP using ordinary Kriging [3].

**3.3. WQP Retrieval Performance Analysis Metrics.** To determine and compare the performance between the sensor-based empirical models in the retrieval of water quality parameter, the regression results were compared with the *in situ* laboratory measurements using the following error matrices: Pearson correlation coefficient  $R$ , coefficient of determination  $R^2$ , mean absolute error (MAE), root mean square error (RMSE), and mean absolute percent error (MAPE%) in Table 4. In Table 4,  $x_i$  and  $y_i$  are, respectively, the laboratory-measured (observed) and the regression model-predicted water quality parameter concentrations at each sample point  $i$  and for the  $n$  samples.

Figure 2 presents a summary of the methodological approach in the retrieval of the water quality parameters.

## 4. Results and Discussions

**4.1. Signal-to-Noise Ratio (SNR) on Clear Water Bodies.** As proposed in Section 3.1.2, the sensors are compared in terms of their responsive SNR indices at specific spectral bandwidths. From spatially uniform and clear water bodies, the Sentinel-2A and Landsat OLI SNR radiometric performances were determined using L1C images for the visible and NIR bands. Notably, for land targets, [33] provided the SNR requirements for the typical radiance. Figure 3 shows the SNR computed from the average of the local window area from the ratio of the area mean to standard deviation ( $\text{SNR} = \mu/\sigma$ ). The SNR is determined using a  $3 \times 3$  pixel kernel on the mean TOA. The selected water body areas are outlined in Figure 4 (spatial areas A and B).

In the three visible bands blue (490 nm), green (560 nm), and red (665 nm), the Landsat OLI bands exhibit SNR values of 2-3 times higher than the corresponding MSI bands. In the coastal aerosol bands (443 nm), MSI is at approximately 30% higher than OLI in SNR. Theoretically therefore, it can be inferred that since SNR is proportional to the square root of the area of a pixel [34], then the aggregated 20 m-30 m spatial resolution of the MSI bands should be able to offer similar or better radiometric quality in the visible bands as compared to the OLI bands for the clear water areas. The same argument can be applied to the aggregated 60 m OLI coastal aerosol band with 30% higher SNR in the corresponding MSI band. With band adjustment and scaling, the overall radiometric quality of MSI and OLI products can be deduced to be comparable over aquatic systems [35]. The multisensor band adjustment and scaling are presented further in the methodology section.

TABLE 4: Water quality estimation performance analysis metrics.

| Error estimator  | Error equation  |
|--|---|
| Pearson correlation coefficient and coefficient of determination | $R = \sum_{i=1}^n (y_i - \bar{y}) \cdot (x_i - \bar{x}) / \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$ , $R^2$ |
| Mean absolute error  | $MAE = 1/n \sum_{i=1}^n  x_i - y_i $  |
| Root mean square error   | $RMSE = \sqrt{1/n \sum_{i=1}^n (x_i - y_i)^2}$  |
| Normalized mean absolute error percent                           | $NMAE\% = \left[ 1/n \sum_{i=1}^n ( x_i - y_i /x_i) \right] * 100$  |

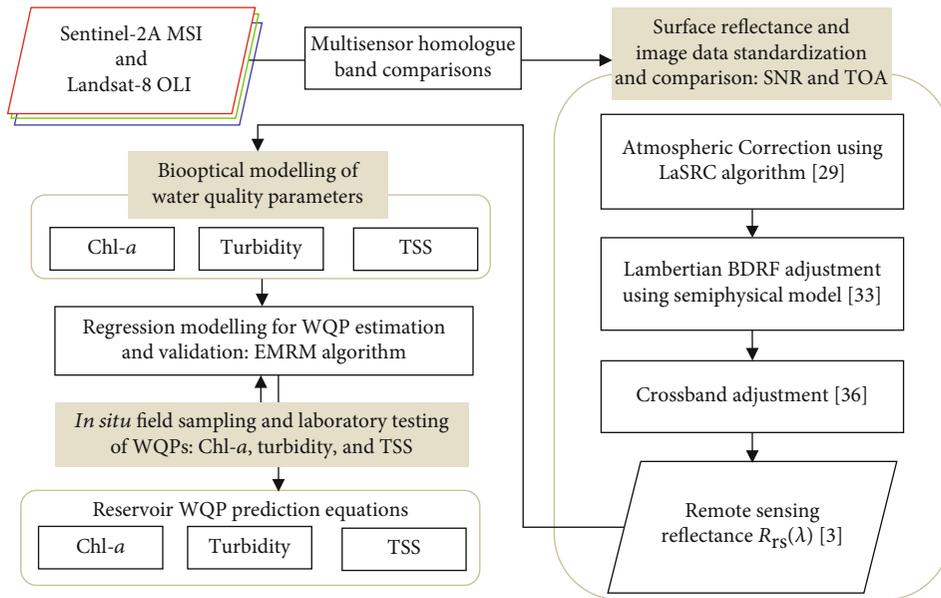


FIGURE 2: Schematic workflow of the approach for the retrieval of water quality parameters from Landsat OLI and Sentinel-2A MSI. WQPs are the water quality parameters, EMRM is empirical multivariate regression modelling algorithmic approach, SNR is the signal-to-noise ratio, and TOA is the top of atmosphere reflectance.

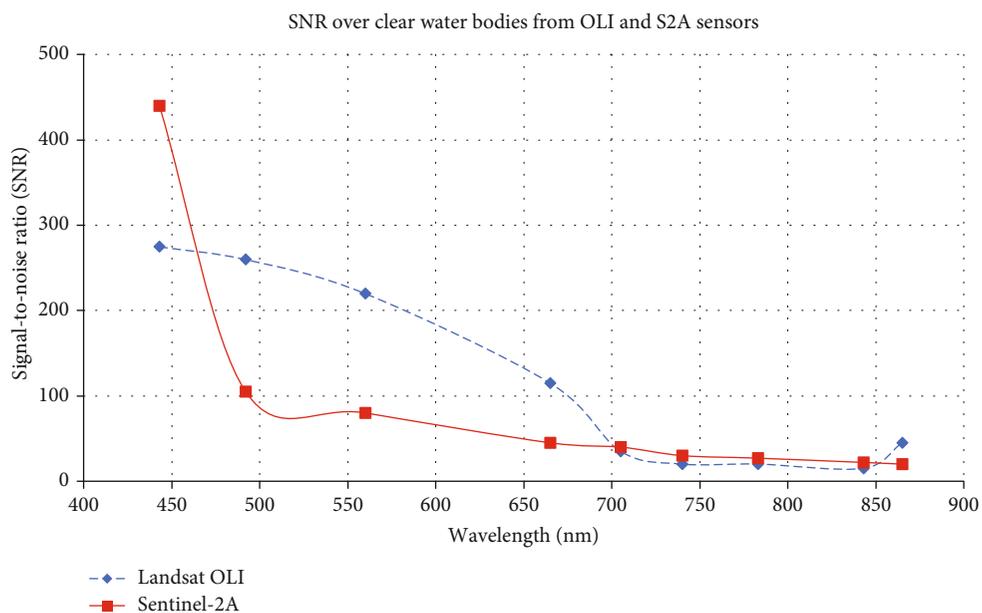


FIGURE 3: SNR on clear water bodies in the visible and NIR bands from Sentinel-2A MSI and Landsat-8 OLI.

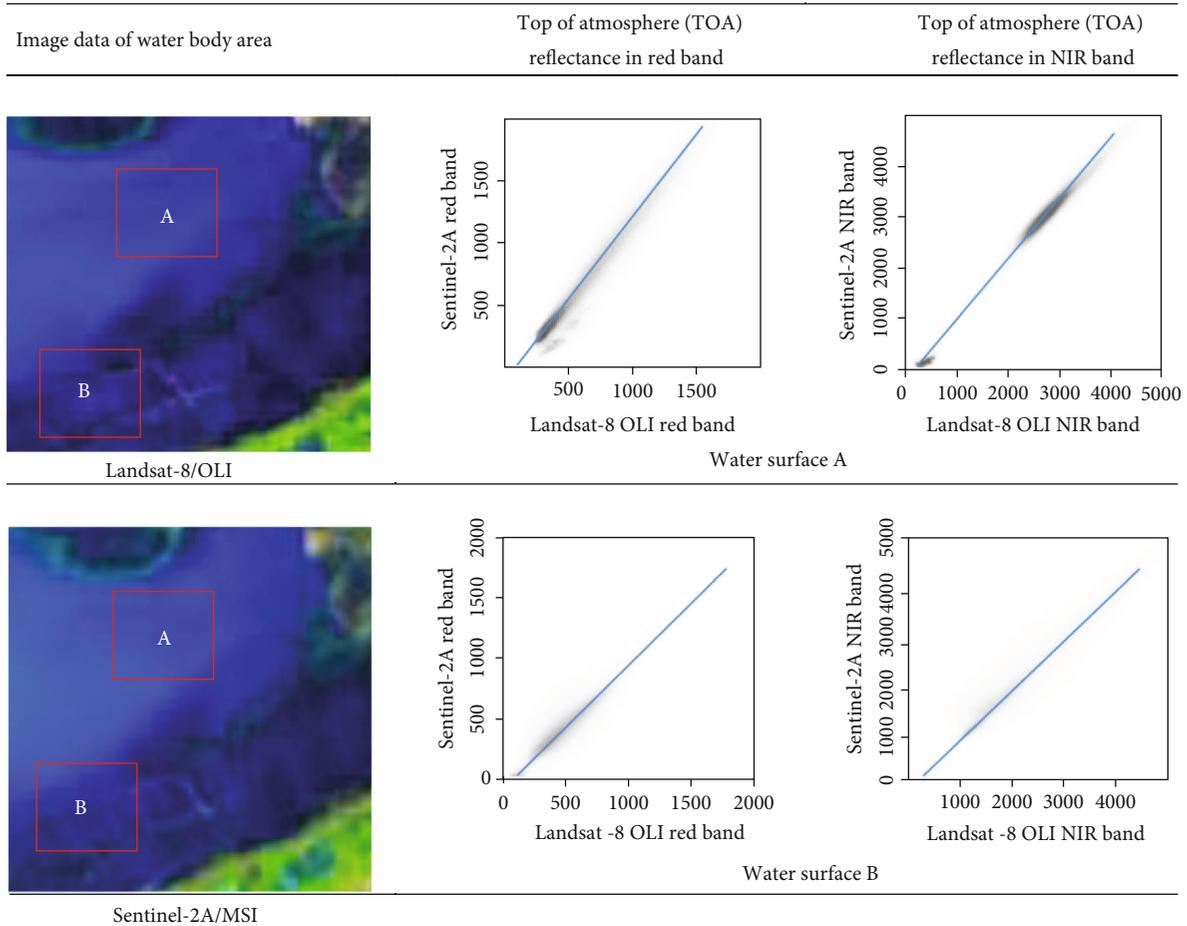


FIGURE 4: The TOA reflectance correlation scatterplots from Landsat OLI and Sentinel-2A reflectance in red and near-infrared bands for clear water surfaces (A and B).

**4.2. Sentinel-2 MSI and Landsat-8 OLI TOA Reflectance Comparison.** When comparing the TOA and surface reflectance from the two sensors, the spatial subsets (A and B) for the selected water bodies are shown in Figure 4. In terms of the interband spectral responses for the selected water body areas, the scatterplots between the Landsat OLI and Sentinel-2A spectral bands showed a high correlation of  $R^2 > 0.876$ , for all the seven homologue bands in Table 1. The regression plots in Figure 4 present sample results from the correlations between the red and NIR bands from the spatial subset of the two water body scenes A and B. The red and NIR bands represent the reflectance in the visible and NIR bands, which are commonly used in water quality modelling [3, 36–38]. The regression results indicate that the two sensors contain nearly similar spectral radiance information, despite the marginal difference in the SNR (Figure 3).

Figure 4 shows that for the same scene water body, the sensors have a close crosscorrelation in the red and NIR bands, which is also observed in the TOA reflectance plot in Figure 5 with the TOA from Landsat OLI wavelengths being slightly higher than that from Sentinel-2A bands. In Figure 5, the TOA reflectance values are scaled by a factor of 10,000 and the 10 m Sentinel-2A bands were resampled to the 30 m spatial resolution to match the Landsat OLI.

The moderately higher TOA reflectance in Landsat OLI confirms the higher SNR. The results in Figure 5 show that the Landsat OLI bands present a slightly higher TOA reflectance than the Sentinel-2A bands.

**4.3. Comparisons of Sensor Remote Sensing Reflectance  $R_{rs}(\lambda)$ .** A comparison of the remote sensing reflectance from the 15 sampling points from the Sentinel-2A and Landsat OLI is presented in Figure 6, indicating that after sensor band adjustment, the reflectance in the sensor bands ranges from 0.003 to 0.009  $\text{Sr}^{-1}$ . For both sensors, the visible and NIR bands show characteristically higher reflectance as compared with the SWIR bands over the reservoir. The sampling stations T2, T6, T9, and T11 depict characteristically higher reflectance than the other sampling points. These sampling points, respectively, represent regions within the dam where there is inflow, the edge of the dam, the deepest and most shallow sections of the dam. In general, the reflectance pattern within the dam follows the same pattern from T1 to T15 especially in Sentinel-2A data, with the deepest point having the highest reflectance of 0.25  $\text{Sr}^{-1}$  in Landsat OLI and 0.087 in Sentinel-2A MSI NIR and blue bands, respectively. In Figure 6, OLI shows the highest reflectance in the visible bands (B2–B4) and NIR (B5) bands, while Sentinel-2A shows relatively higher

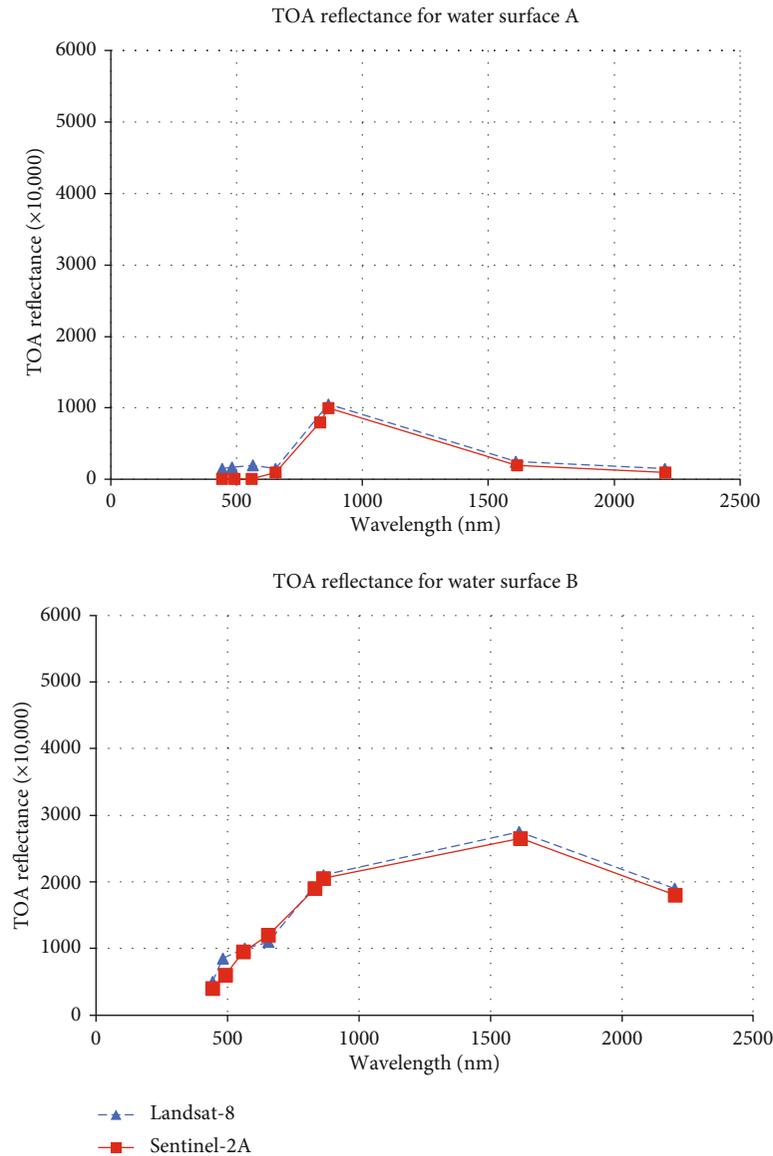


FIGURE 5: Comparison between TOA reflectance from Landsat-8 OLI and Sentinel-2A MSI for clear water surfaces.

TOA reflectance in all the bands at above  $0.003 \text{ Sr}^{-1}$ . This observation is similar to the observed SNR results in Figure 3, and the near-uniform reflectance from the two sensors can be attributed to the band adjustment using the LaSRC algorithm [27].

#### 4.4. Water Quality Estimations from Sentinel-2A and Landsat OLI

**4.4.1. Retrieval of Chlorophyll-*a*.** The *in situ* test results determined the concentrations of chlorophyll-*a* to vary considerably within the reservoir with values between  $6.23 \mu\text{g/L}$  and  $113 \mu\text{g/L}$  and an average of  $29.91 \mu\text{g/L}$ . Since the reservoir is protected from public use, the observed Chl-*a* concentrations are attributed to the inflow of fertilizer leachate into the reservoir from the surrounding farms. The leachate process carries nutrients which encourage the growth of algae in

the reservoir. This is likely to cause severe eutrophication if not monitored on a continual basis.

From the empirical regression modelling, Figures 7 and 8, respectively, present the best regression models for the estimation of the concentration of Chl-*a* from Sentinel-2A and Landsat OLI. For Sentinel-2A (Figure 7), a second-order polynomial fit was found to be suitable for the retrieval of Chl-*a* using  $R_{rs}(\lambda)$  from the difference between the green (B3) and the SWIR-1 (B11) band, with  $R^2 = 0.7015$ . This was approximately 15% lower in accuracy as compared to the linear model derived from the prediction of Chl-*a* from Landsat OLI with  $R^2 = 0.8581$  from the green band (B3) (Figure 8). While the green band is observed to be significant in the detection of Chl-*a* from both sensors, the advantage of Landsat OLI could be because of its wider spectral bandwidth in the green band, i.e., approximately 57 nm as compared to 35 nm in Sentinel-2A (Table 2),

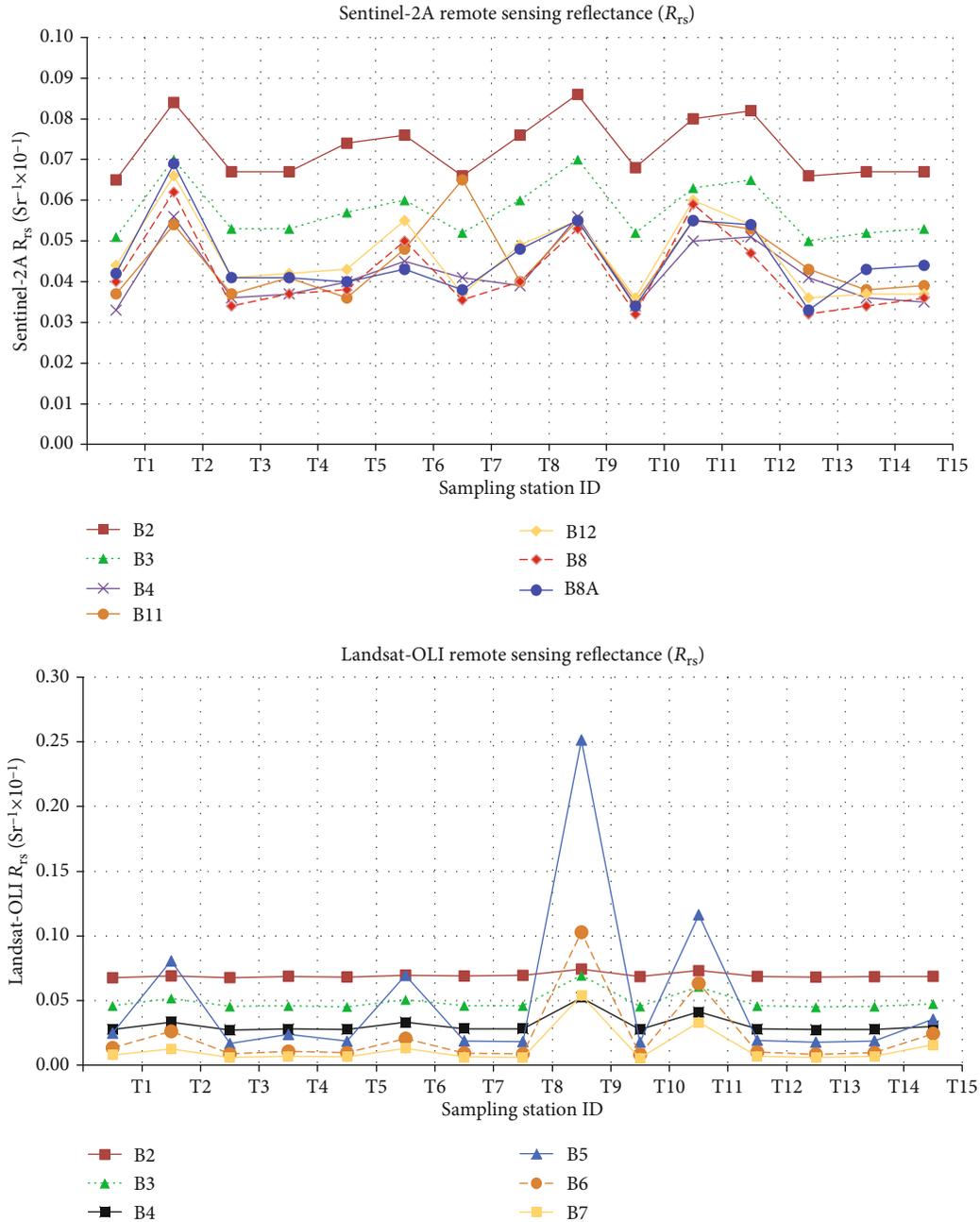


FIGURE 6: Remote sensing reflectance variations over sampling points.

and therefore, the placement of the green band and the SNR in the Landsat OLI is suitable for Chl-*a* retrieval.

The estimated Chl-*a* from Sentinel-2 and Landsat-8 both have a lower RMSE value compared to the mean *in situ*-measured Chl-*a* (29.91  $\mu\text{g/L}$ ). This confirms the appropriateness of the developed regression model for estimating Chl-*a* in the case study dam at the given acquisition.

A summary of the performance of the two sensors is given in Tables 5 and 6 for the best five fit EMRM results. The results show that Landsat OLI is able to detect the presence of Chl-*a* with  $R^2$  of up to 0.86, as compared to Sentinel-2A with  $R^2$  of up to 0.70. Similarly, the statistical trends from MAE, NMAE%, and RMSE are the same at  $p$  value = 0.055,

with lower MAE and RMSE in Landsat OLI, but relatively the same NMAE% error magnitudes for Chl-*a* retrieval from both sensors. The inference from the results is that both the sensors are suitable for the retrieval of Chl-*a*, specifically using the green and blue bands, and thus, as opposed to the results from [39], the red and NIR bands are not informative in the retrieval of water quality in the inland waters in this case study. The green and blue wavelengths are suitable for the detection of algal bloom due to the high reflectance of green algal matter, which is in the bluish color waters.

Compared to previous study results on the estimation of Chl-*a*, [38] compared Landsat-8 L1T and Sentinel-2A L1C with *in situ* measurements for chlorophyll-*a* predictions in

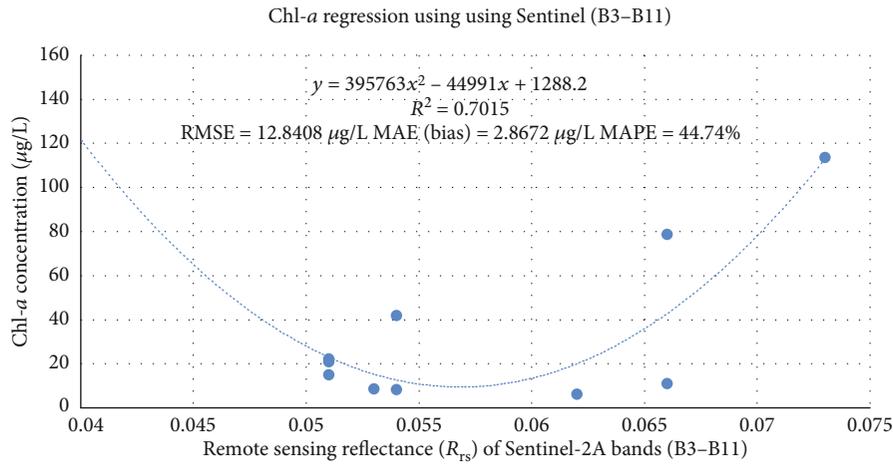


FIGURE 7: *In situ*-derived chlorophyll-*a* and Sentinel-2A band reflectance.

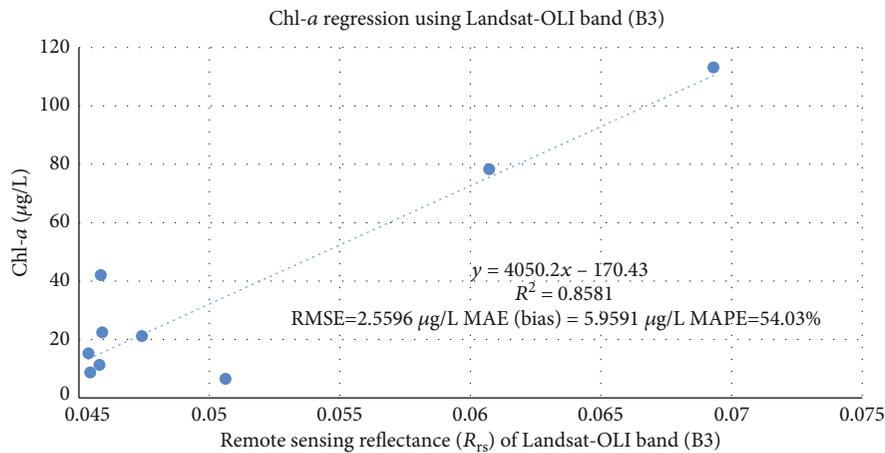


FIGURE 8: Landsat band reflectance and *in situ* Chl-*a* regression.

TABLE 5: Regression model for the retrieval of chlorophyll-*a* from Sentinel-2A bands.

| Sentinel-2A regression model equation for Chl- <i>a</i> estimation | Band combination for deriving $R_{rs}(=x)$ | $R^2$  | MAE     | NMAE%   | RMSE    |
|--|--|--------|---------|---------|---------|
| $y = 466683x^2 - 39497x + 835.87$                                  | B3 + (B12-B11)                             | 0.6915 | 0.1613  | 8.7155  | 3.0357  |
| $y = 378524.04x^2 - 30345.60x + 616.37$                            | B3   | 0.6300 | 3.4864  | 73.1354 | 27.3437 |
| $y = 395763x^2 - 44991x + 1288.2$                                  | B2-B11                                     | 0.7015 | 2.8672  | 44.7386 | 12.8408 |
| $y = 246622x^2 - 38346x + 1497.6$                                  | (B2 + B12)/2                               | 0.6291 | 6.0437  | 67.4678 | 23.2026 |
| $y = 1748431.97x^2 - 126132.79x + 2289.87$                         | (B2-B11)/2                                 | 0.6500 | 10.2352 | 73.3885 | 18.6211 |

TABLE 6: Regression models for the estimation of chlorophyll-*a* from Landsat OLI bands.

| Landsat-8 OLI regression model equation for Chl- <i>a</i> estimation | Band combination for deriving $R_{rs}(=x)$ | $R^2$  | MAE    | NMSE%   | RMSE   |
|--|--|--------|--------|---------|--------|
| $y = 3188.9x - 349.09$   | B2 + B3                                    | 0.8505 | 1.4128 | 59.5924 | 8.2940 |
| $y = 4050.2x - 170.43$   | B3   | 0.8581 | 2.5596 | 54.0317 | 5.9591 |
| $y = 2749.1x - 539.88$   | B2 + B3 + B5                               | 0.8568 | 0.4808 | 60.9354 | 9.9056 |
| $y = 402.73 \ln(x) + 888.88$   | B2 + B3                                    | 0.8522 | 1.9835 | 60.1054 | 7.7013 |
| $y = 593.38 \ln(x) + 964.48$   | B2 + B3 + B4                               | 0.8507 | 0.8458 | 61.4433 | 9.4775 |

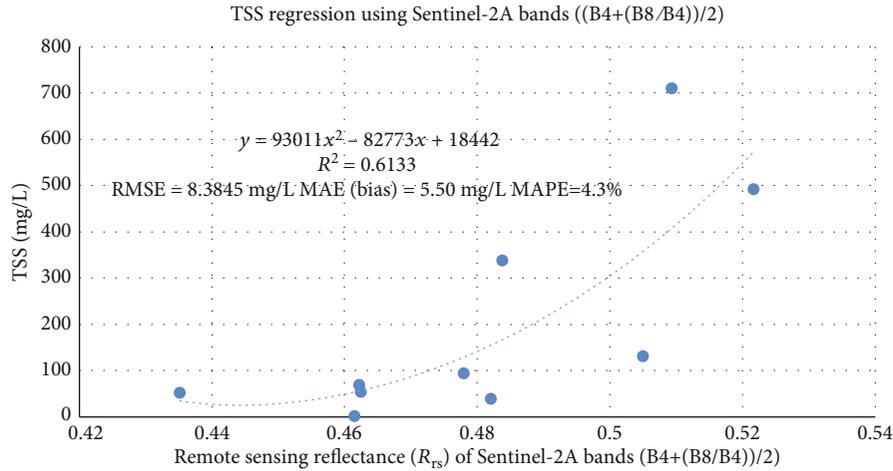


FIGURE 9: Measured TSS plotted against Sentinel-2A reflectance.

the Cassaffouth Reservoir in Argentina. For Landsat, the model with the combination of the blue band and blue/NIR and green/NIR band ratios produced the best results with  $R^2 > 0.89$ . The best Sentinel MSI model was a two-variable model using both the blue band and the (green/NIR band ratio with  $R^2 = 0.880$ . [13] also estimated chlorophyll-*a* for Lake Peipsi and Lake Vörtsjärvi using Sentinel MSI with  $R^2 = 0.83$ , using blue and green bands. In attempts to estimate chlorophyll-*a*, [40] showed that Landsat OLI produced average but better results for chlorophyll-*a* estimations for the case study of Lake Trichonis in Greece, as compared to its predecessors in the Landsat series, using the visible and NIR bands. [41] investigated the potential of deriving chlorophyll-*a* and turbidity using Secchi disk depth in six (6) natural and five (5) artificial lakes in Greece, using Landsat-8's blue, red, and SWIR-2 bands. The regression validation indicated successful correlations with  $R^2$  of 0.75 for chlorophyll-*a* estimations. [11, 39] derived the chlorophyll-*a* for Barra Bonita tropical reservoir using Landsat OLI. The results showed that the NIR-red, NIR-green, and NIR-blue band ratios yielded the best results with  $R^2$  greater than 0.70 for the estimation of Chl-*a* concentrations.

In conclusion and compared to previous studies, the current study further confirms the suitability of the visible and NIR bands in the prediction of chlorophyll-*a* in inland water bodies. The detection of the concentration of chlorophyll-*a* using the developed empirical equations can be an indicator for monitoring the abundance of algae in inland water bodies, since it presents active optical properties in the visible and near-infrared wavelengths.

**4.4.2. Estimation of TSS Concentration.** The concentrations of suspended particulate matter from the laboratory measurements were low and within acceptable limits ranging between 39 mg/L and 711 mg/L, with a reservoir average of 187 mg/L. The highest concentrations were in the same regions as for Chl-*a* and are attributed to the inflow of water from a river at an area of low depth which leads to the agitation of the sediments from the bottom of the reservoir.

Because the reservoir is deep, the sediments tend to settle faster leading to generally low concentrations of particulate matter. However, the high presence of particulate matter at a given time and region propagates the concentration of algal bloom and hence the observed high concentrations of Chl-*a* in the same reservoir region.

Using the EMRM algorithmic approach, the best results for the estimation of TSS concentration from Sentinel-2A were obtained using a second-order polynomial function relating the red band (B4) and the NIR band (B8) with  $R^2$  of 0.6113 (Figure 9), as compared to  $R^2 = 0.9249$  from the Landsat OLI band ratio between the green band (B3) and the blue band (B2) using linear regression modelling (Figure 10). While the mean concentration of TSS from the in situ measurements is at 187 mg/L, the RMSE and bias are averaged and less than 10 mg/L. This confirms the validity of the developed model in estimating TSS within the reservoir.

The overall performance of the satellite sensors in the retrieval of TSS within the reservoir is summarized in Tables 7 and 8 for the best five results. As in the retrieval of Chl-*a*, the visible and NIR bands are the optimal bands in the estimation of TSS. The accuracy of estimation of TSS from Landsat OLI (Table 8) is notably higher than that from Sentinel-2A by at least 30% as measured in terms of  $R^2$ . The MAE, NMAE%, and RMSE are also observed to follow the  $R^2$  trend of higher values from Sentinel-2A predictions as compared to the predictions from Landsat OLI. The results in Tables 7 and 8 show that the placement of the visible bands in Landsat is significant in TSS estimation.

The significance of OLI band 3 and band 2 in TSS estimation is noted and is attributed to the fact that the deeper the water, the more blue light it reflects and hence the higher the reflectance in blue and green bands. It is also known that the particle settlement is higher in deep quiescent water. Therefore, by relation, the higher the blue light detected, the lower the concentration of particles in the water. In the case of the green band, most of the suspended matter in the water was in the form of algae. Therefore, the higher the

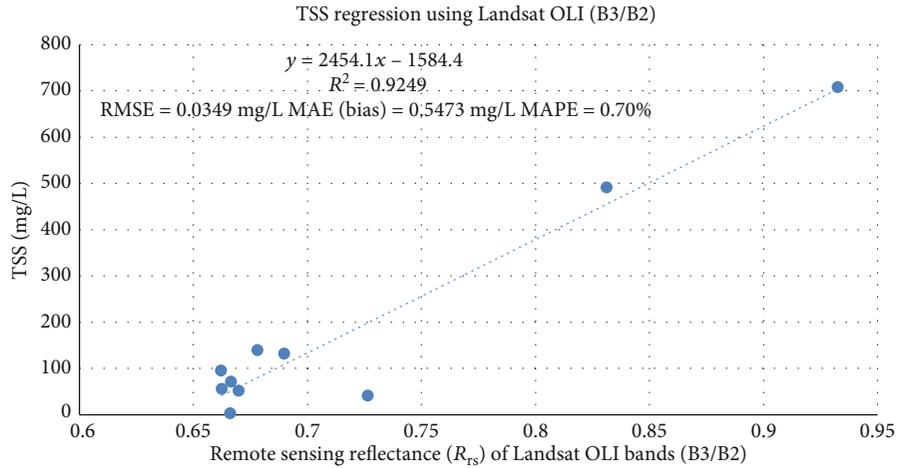
FIGURE 10: Landsat band remote sensing reflectance and *in situ*-measured TSS concentration.

TABLE 7: Regression model for the estimation of TSS from Sentinel-2A MSI.

| Sentinel-2A regression model equation for estimation of TSS | Band combination for deriving $R_{rs}$ ( $= x$ ) | $R^2$  | MAE     | NMAE%   | RMSE    |
|---|--|--------|---------|---------|---------|
| $y = 93011x^2 - 82773x + 18442$                             | $(B4 + (B8/B4))/2$                               | 0.6133 | 5.4969  | 4.3039  | 8.3845  |
| $y = 89013x^2 - 78936x + 17526$                             | $(B8A + (B8/B4))/2$                              | 0.5940 | 9.5691  | 6.8105  | 13.4793 |
| $y = 83708x^2 - 73811x + 16299$                             | $(B8 + (B8/B4))/2$                               | 0.5998 | 6.6792  | 7.5058  | 12.9244 |
| $y = 19420x^2 - 26272x + 8945.4$                            | B4/B3  | 0.4846 | 23.7299 | 15.7569 | 24.8370 |
| $y = 46042x^2 - 32585x + 5822.6$                            | $(B3 + (B4/B3))/2$                               | 0.4722 | 13.9118 | 12.9250 | 21.7248 |

TABLE 8: Regression model for the estimation of TSS from Landsat-8 OLI.

| Landsat OLI regression model equation for estimation of TSS | Band combination for deriving $R_{rs}$ ( $= x$ ) | $R^2$  | MAE    | NMAE%   | RMSE    |
|---|--|--------|--------|---------|---------|
| $y = 2781.9x - 1393.5$                                      | B2   | 0.9240 | 4.7391 | 12.3842 | 3.3185  |
| $y = 27176x - 1184.6$                                       | B3   | 0.9245 | 6.8510 | 29.896  | 12.8425 |
| $y = 2454.1x - 1584.4$                                      | B3/B2  | 0.9249 | 0.5473 | 0.7000  | 0.0349  |
| $y = 1714.7 \ln(x) + 1171$                                  | B3/B4  | 0.9161 | 2.3885 | 9.3736  | 2.1329  |
| $y = 1905.3 \ln(x) + 821.01$                                | B3/B2  | 0.9159 | 3.7496 | 12.1181 | 4.5714  |

amount of light reflected in the green wavelength, the higher the concentration of algae and thus the higher the amount of suspended matter.

Compared to previous studies on TSS estimations using satellite data, [36] utilized Landsat OLI for retrieval of suspended sediment concentrations in the Orinoco River in Venezuela. The best performance for TSS estimations was from the red and near-infrared wavelengths, which is similar to the results from Sentinel-2 MSI in this study. [21] also found that for suspended particulate matter predictions for Poyang Lake in China using empirical models, Sentinel-2A MSI band 7 was suitable for sediment-laden waters while Sentinel-2A MSI band 4 was suitable for clear waters. In Playa Colorada Bay, Mexico, [42] generated empirical models for TSS and turbidity estimations using the Landsat imagery. The correlation results between the Landsat reflectance

and the *in situ* measurements indicated that both water quality parameters could be estimated with  $R^2$  of between 0.637 and 0.955, which is comparable with current results using Landsat OLI. [37] also estimated TSS and turbidity for Wular Lake in Kashmir. The parameters were correlated with single Landsat OLI bands and produced  $R^2$  greater than 0.5, while [43] estimated TSS with  $R^2$  of 0.709 using visible bands for Poteran Island waters in East Java, Indonesia.

The results from the current study show the significance of the green, red, and near-infrared wavelengths in the estimations of TSS in inland reservoir water bodies, with specially higher accuracy using Landsat-8 OLI wavelengths. For TSS retrieval from Sentinel-2 MSI, however, the marginal performance could be attributed to interactions of the optical electromagnetic radiations with the atmosphere and resulting low water surface reflectance into the sensor (glint). Thus,

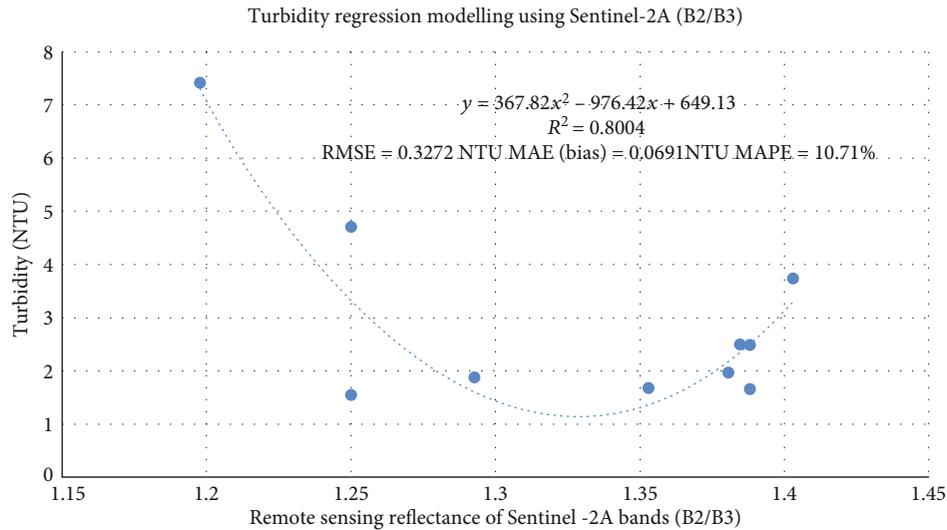


FIGURE 11: Laboratory-measured turbidity and Sentinel-2A band reflectance.

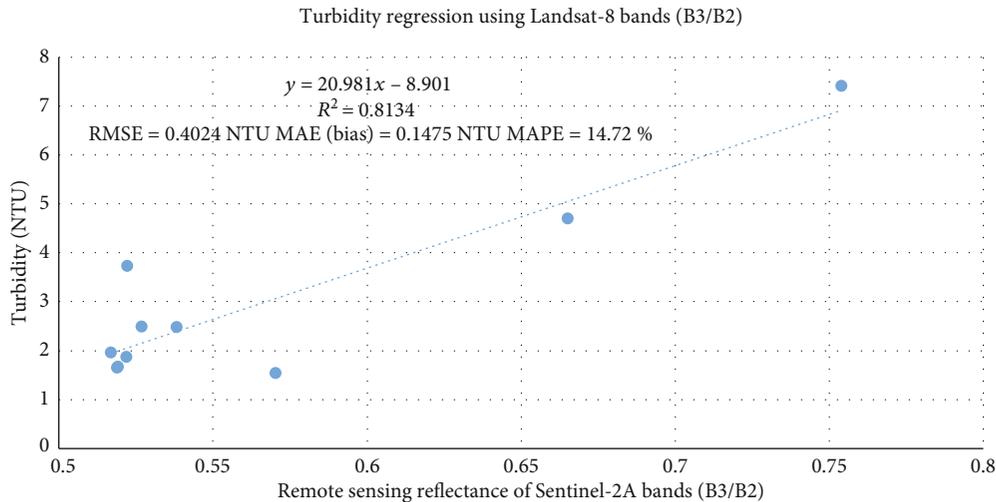


FIGURE 12: Landsat band reflectance and *in situ* turbidity measurements.

further correction of the Sentinel MSI  $R_{rs}$  signal is proposed to improve the accuracy in modelling TSS using Sentinel-2 data [13, 38].

**4.4.3. Estimation of Turbidity Concentration in Reservoir Waters.** With *in situ* turbidity varying from 1.36 NTU to 7.42 NTU and averaging at 2.83 NTU, the reservoir is characterized by low turbid waters. The low turbidities could be attributed to low flows into the reservoir, especially during the period in which the water samples were collected. With minimum inflow of sediment-laden rainwater discharge and the slow velocities of the river water, the concentrations of sediments that would cause reservoir turbidity were low. Further, at a maximum depth of approximately 150 m, the reservoir depth enables the settlement of sediments at the bottom of the reservoir with minimal potential of resuspension by water currents and waves. The higher degree of settling of sediments therefore leads to low turbidity in the reservoir.

From the empirical multivariate regression modelling comparing the satellite reflectance data and the laboratory-measured water quality factor, the study results show that the turbidity was estimated from the two sensors with the same accuracy of  $R^2 > 0.7$  (Figures 11 and 12). For both the sensors, the visible bands (blue, green, and red) are observed to be predominant in the retrieval of turbidity within the reservoir. In Figure 11, the best estimate of turbidity using Sentinel-2A is at  $R^2 = 0.8004$ , which is comparable to turbidity estimates using Landsat OLI with  $R^2 = 0.8134$ . The models accurately predicted the concentrations of turbidity from both sensors as the RMSE and bias error measures were less than the minimum and average *in situ* turbidity.

Tables 9 and 10 give a summary of the best regression models for the retrieval of turbidity, showing significant contributions from the visible bands from both sensors. The error metrics MAE, NMAE%, and RMSE are, respectively, in the same magnitude range for both sensors. Like in the

TABLE 9: Regression model for the estimation of turbidity from the Sentinel-2A MSI sensor.

| Sentinel-2A regression model equation for estimation of turbidity | Band combination for deriving $R_{rs}$ ( $= x$ ) | $R^2$  | MAE    | NMAE%    | RMSE   |
|---|--|--------|--------|----------|--------|
| $y = 367.82x^2 - 976.42x + 649.13$                                | B2/B3  | 0.8004 | 0.0691 | 10.7144  | 0.3272 |
| $y = 971.47x^2 - 1468x + 555.84$                                  | B3/B2  | 0.7990 | 0.0746 | 10.4636  | 0.3290 |
| $y = 725.32x^2 - 858.52x + 255.91$                                | B4/B3  | 0.7353 | 0.7609 | 47.9432  | 1.5866 |
| $y = 118.8x^2 - 401.92x + 341.62$                                 | B2/B4  | 0.7372 | 0.7656 | 45.71415 | 1.5877 |
| $y = 387.41x^2 - 1103x + 786.37$                                  | B1 + (B1/B2)                                     | 0.7592 | 0.0035 | 6.9895   | 0.2192 |

TABLE 10: Regression model for the estimation of turbidity from the Landsat-8 OLI sensor.

| Sentinel-2A regression model equation for estimation of turbidity | Band combination for deriving $R_{rs}$ ( $= x$ ) | $R^2$  | MAE    | NMAE%   | RMSE   |
|---|--|--------|--------|---------|--------|
| $y = 20.981x - 8.901$   | B3/B2  | 0.8134 | 0.1475 | 14.7203 | 0.4024 |
| $y = 102.56x - 5.5003$  | B3 + B4  | 0.8125 | 0.1066 | 9.4289  | 0.7341 |
| $y = 90.319x - 10.775$  | B2 + B3 + B4                                     | 0.8123 | 0.1175 | 9.9483  | 0.7836 |
| $y = 20.254 \ln(x) + 46.009$                                      | B2 + B3  | 0.7978 | 0.0489 | 14.5470 | 0.6069 |
| $y = 14.735 \ln(x) + 30.802$                                      | B2 + B3 + B4                                     | 0.7954 | 0.0865 | 10.5043 | 0.7350 |

TABLE 11: Descriptive statistics of the laboratory-measured and predicted water quality parameters (SD = standard deviation; CV = coefficient of variation; SE = standard error).

| Water quality parameter and estimation and data source | Sample ( $n$ )                  | Min | Max   | Median | Average | SD     | CV (%) | SE     |       |
|--|---------------------------------|-----|-------|--------|---------|--------|--------|--------|-------|
| Chl- $a$ ( $\mu\text{g/L}$ )                           | <i>In situ</i> WQP measurements | 15  | 6.24  | 113.73 | 15.16   | 29.91  | 31.65  | 105.82 | 8.17  |
|  | Landsat-8 OLI                   | 15  | 13.29 | 110.26 | 15.44   | 30.77  | 28.42  | 92.39  | 7.34  |
|  | Sentinel-2A MSI                 | 15  | 11.38 | 221.80 | 23.04   | 40.82  | 56.19  | 137.65 | 14.51 |
| Turbidity (NTU)  | <i>In situ</i> WQP measurements | 15  | 1.36  | 7.42   | 1.97    | 2.83   | 1.69   | 59.64  | 0.42  |
|  | Landsat-8 OLI                   | 15  | 1.94  | 6.92   | 2.05    | 2.88   | 1.46   | 50.77  | 0.38  |
|  | Sentinel-2A MSI                 | 15  | 1.23  | 26.25  | 2.48    | 4.99   | 6.51   | 130.52 | 1.68  |
| TSS (mg/L)   | <i>In situ</i> WQP measurements | 15  | 39.00 | 711.00 | 95.00   | 187.00 | 197.88 | 102.23 | 49.36 |
|  | Landsat-8 OLI                   | 15  | 38.12 | 702.60 | 60.23   | 169.68 | 195.70 | 115.34 | 50.53 |
|  | Sentinel-2A MSI                 | 15  | 35.51 | 573.74 | 127.78  | 177.06 | 157.98 | 89.22  | 40.79 |

estimation of Chl- $a$  and TSS, the linear and polynomial regression models are the best for the retrieval of the WQP. The turbidity estimation results are particularly important since turbidity is considered one of the most significant water quality parameters [44].

In the estimation of turbidity, the visible and NIR bands are observed to yield the best results. Comparatively, the accuracy of turbidity estimations is nearly equal for the two optical sensors indicating that the reflectance of suspended particulates is much higher in the lower wavelengths, irrespective of the band and or the bandwidth.

For the assessment of the extraction of turbidity as the significant water quality parameter that relates directly to the underwater light penetration that affects productivity, [44] showed from several band combinations that Landsat OLI band 4 (red) has the best correlation with the *in situ*-measured turbidity with a coefficient of determination  $R^2$  of 0.84, scatter index of 0.22, and RMSE of 0.28 NTU. [45] also

recently demonstrated using the multivariate regression model and gene expression programming (GEP) that reservoir turbidity could be extracted with  $R^2$  of greater than 90%, using the combination of visible and NIR bands. While high accuracy has been observed in turbidity estimation, in all the studies, atmospheric correction was found to be a factor in the accuracy of the WQP modelling [13, 22, 38].

**4.5. Validation of Water Quality Prediction with In Situ Measurements.** The validation of the developed regression models in Section 4.4 was carried out using the five (5) sampling stations (T2, T5, T7, T8, and T13). The validation results are presented in Table 11, including the statistics from the stations which were used in the model calibration. TSS is observed to have the highest variation in concentration, followed by Chl- $a$  and turbidity (least). From the SD, CV, and SE metrics, Landsat OLI tended to underestimate the concentrations of the water quality

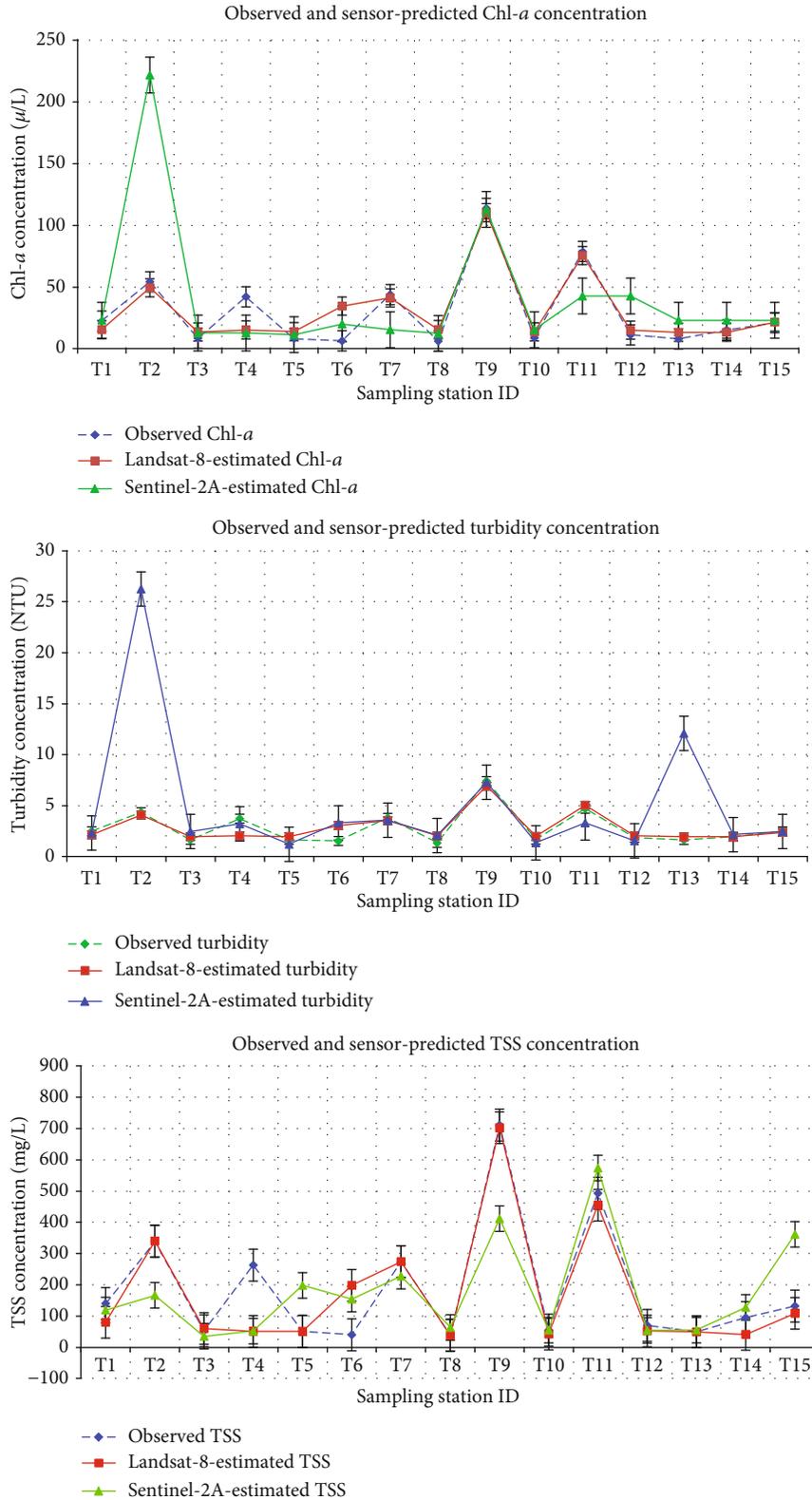


FIGURE 13: Prediction and validation sampling point variations of Chl-*a*, turbidity, and TSS within the reservoir.

parameters, while Sentinel-2A-based models tended to overestimate the same WQP, thus predicting the parameters with a higher coefficient of variation.

The graphical assessment of the validation results in Figure 13 for Chl-*a*, turbidity, and TSS shows that for the five (5) validation sampling stations, the retrieval of Chl-*a* and

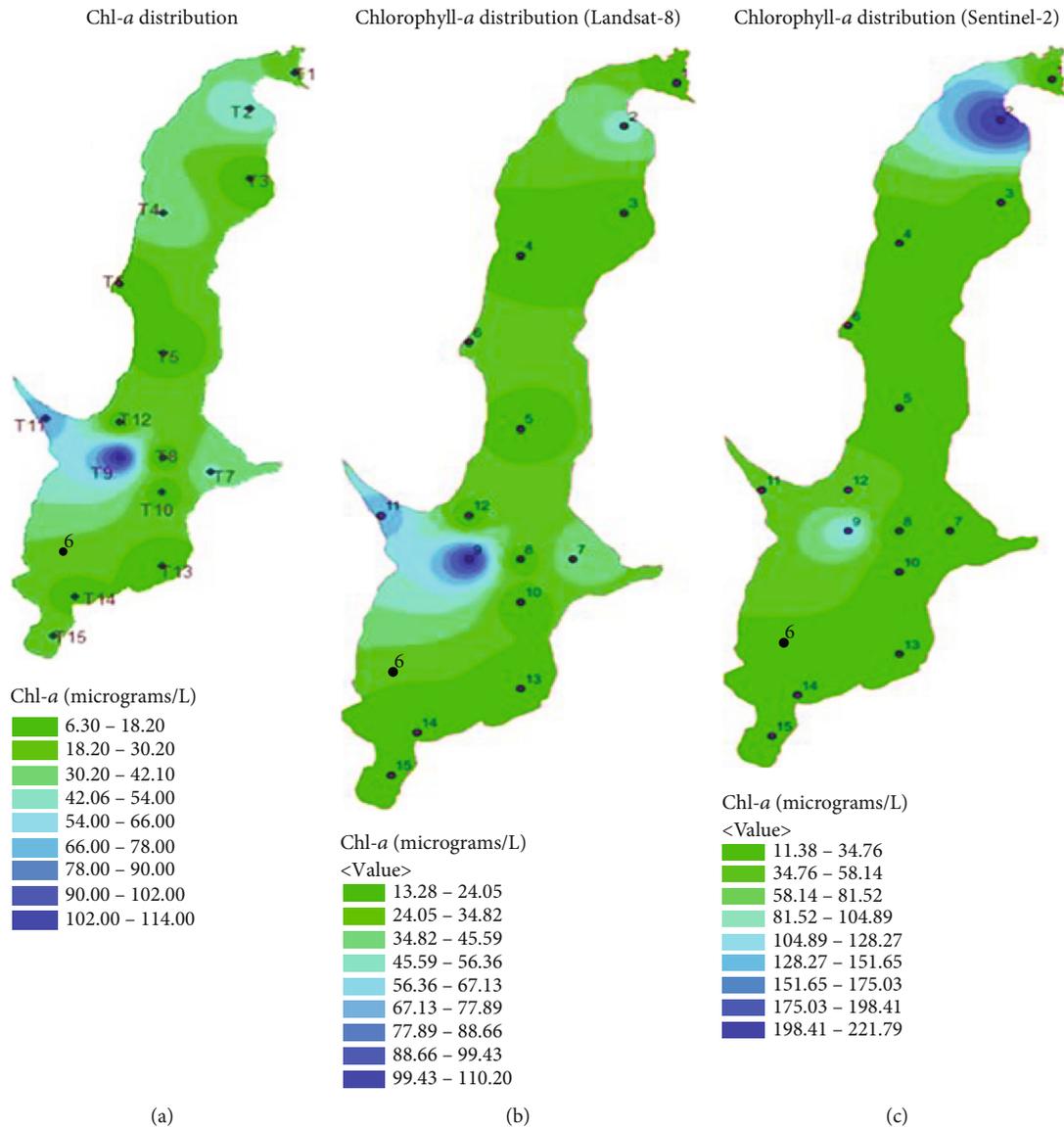


FIGURE 14: Spatial distribution of Chl-*a* concentration from (a) *in situ* measurements and from (b) Landsat OLI and (c) Sentinel-2A.

turbidity matched the observed measurements for both Sentinel-2A and Landsat OLI, except for sampling station T2, where Sentinel-2A grossly overestimated the Chl-*a* and turbidity concentrations by approximately 4 times. For TSS estimation, Table 9 and Figure 13 show lower coincidence between the measured and validated results with standard error at above 40% for both sensors.

The marginally lower performance of the Sentinel-2A in the prediction of the selected WQP may be attributed to the two factors, namely: (i) the sensor's low spectral resolution in terms of the bandwidths and (ii) the lower signal-to-noise ratio (SNR) as compared to Landsat OLI visible and NIR bands. However, in the four visible channels and the NIR bands, the results of the study show that with adequate radiometric corrections or band adjustments, the OLI and MSI are both capable of monitoring the optically active water quality indicators. In synergy with the Landsat-9, the multi-sensor integration of the Sentinel-2A MSI and Landsat-8 will

result in a temporal revisit time of 2–3 days, making it possible to accurately capture the dynamics of the inland water body characteristics for continuous assessments.

In summary, the results for the prediction of chlorophyll-*a*, TSS, and turbidity demonstrate the potential for cost-effective, large-scale, and high-frequency use of optical satellite remote sensing reflectance data in monitoring optically active water elements. However, it is observed that the effectiveness and accuracy of water quality element determination are dependent on the satellite sensor used, the methodology followed, and also the nature of the waters studied.

#### 4.6. Spatial Distribution of Optically Active Water Quality Parameters

**4.6.1. Chlorophyll-*a* Distribution.** Using ordinary Kriging as suggested in [3], the spatial interpolation of the concentrations of Chl-*a* observed and estimated from Sentinel-2A

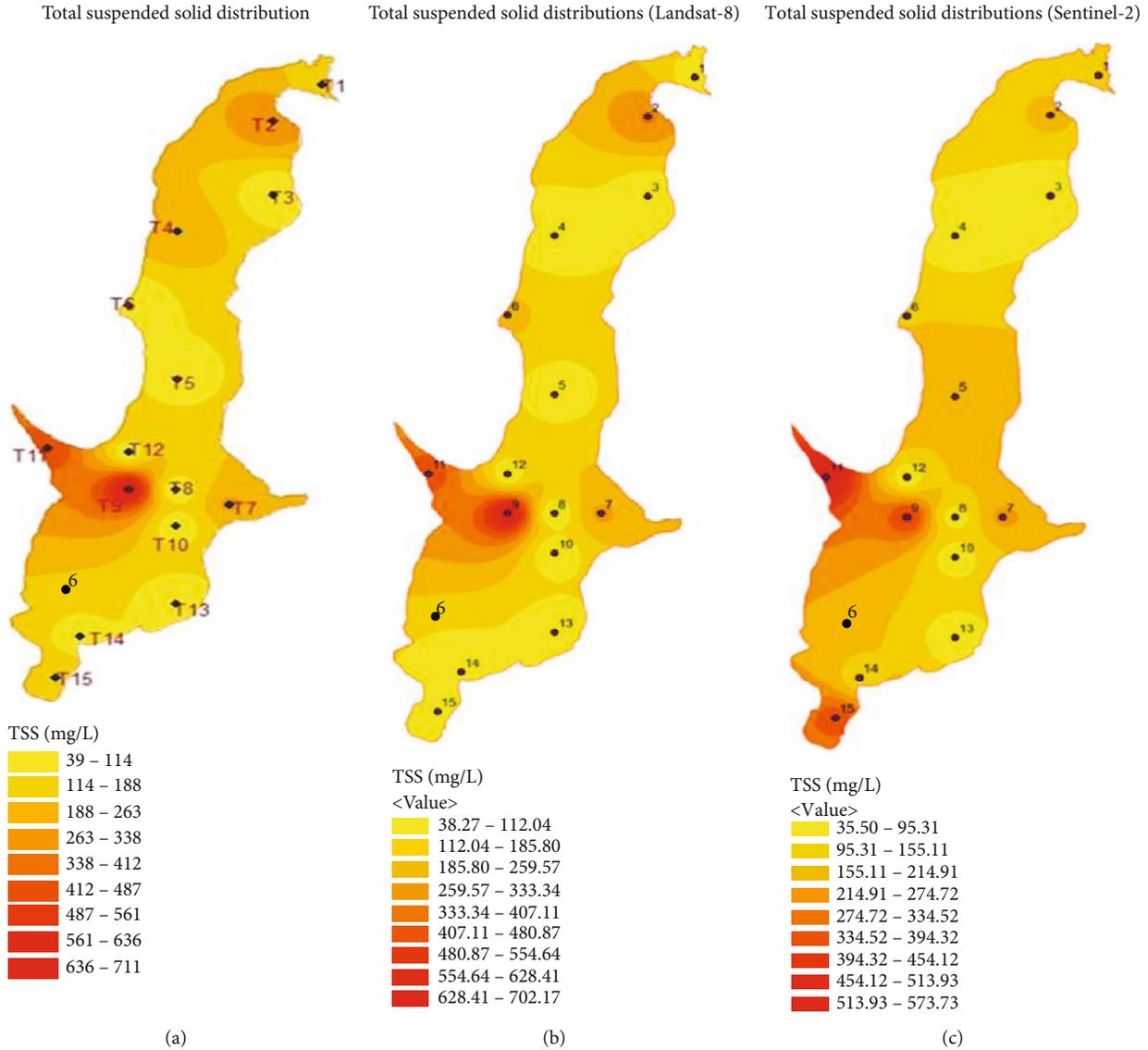


FIGURE 15: Spatial distribution of TSS concentration from (a) *in situ* measurements and (b) Landsat OLI and (c) Sentinel-2A.

and Landsat OLI datasets is presented in Figure 14. It is observed that the chlorophyll distribution trend from the *in situ* measurements and Landsat OLI (Figures 14(a) and 14(b), respectively) is closely correlated with a good coincidence between the high and low chlorophyll-*a* concentration regions. The highest concentrations are observed around the deepest region of the dam at T9, with Landsat-8 OLI slightly underestimating the chlorophyll-*a* concentration (99.43-110.20  $\mu\text{g/L}$ ), as compared to the laboratory-measured concentrations ranging from 102.00 to 114.00  $\mu\text{g/L}$ . On the contrary, for high Chl-*a* concentration regions, Sentinel-2A retrieval does not coincide with that measured and estimated using Landsat OLI as evident in Figure 14(c) with gross overestimation and shift in the spatial location. This difference provides satisfactory evidence as shown in Table 11 and Figure 13, whereby at T2 and T12-T14, the estimation errors of Chl-*a* from Sentinel-2A are marginally higher.

The predominately green band-based Chl-*a* algorithms derived in this study are considered to be suitable for inland

water body systems whereby Chl-*a* is the dominant absorber. For optically complex waters like rivers and polluted lakes, combinations of other water components may result in false Chl-*a* derivations. This can be the case in terrestrially influenced water bodies, where CDOM and nonalgal particles are able to absorb shorter wavelengths in the visible spectrum, thus changing the green reflectance value which may lead to the overestimation of Chl-*a* [46]. For the establishment of the differences between water body systems, the uncertainties and absolute accuracy of the developed algorithms should be considered for Chl-*a*, TSS, and turbidity.

**4.6.2. Distribution of TSS Concentration.** The results in Figure 15 show that the distribution of TSS in the dam was fairly similar between the *in situ* laboratory measurements and the estimations by the Landsat OLI and Sentinel-2A sensors. In the case of TSS and as observed for Chl-*a*, Sentinel-2A tended to slightly overestimate the TSS concentration; however, the peak locations of high and low concentrations



## 5. Conclusions

This study compared Landsat-8 OLI and Sentinel-2A MSI satellite sensors for the retrieval of chlorophyll-*a*, TSS, and turbidity in an inland water reservoir. For clear water surfaces, Landsat-8 exhibited a higher signal-to-noise ratio (SNR) in the visible bands and marginally higher water surface reflectance in the visible and NIR bands as compared to the corresponding Sentinel-2A bands. Using the visible and NIR bands with the empirical multivariate regression model, the presence of Chl-*a* was retrieved from Sentinel-2A with accuracy of 62-70%, which was 15-25% less than the Chl-*a* estimations from Landsat-8. Using the same multi-spectral bands, TSS was estimated from Sentinel-2A with accuracy of between 47 and 62% as compared to Landsat-8's accuracy of more than 90%. For turbidity estimations, the performance of both sensors was nearly equal with accuracy of more than 75%. The results show the potentials of using both sensors in reservoir water quality retrieval, with Landsat-8 performing better in estimating the tested parameters. To improve the accuracy of water quality estimations in inland water bodies, the study recommends the radiometric calibration and synergistic use of Landsat-8/9 and Sentinel-2A/2B for frequent water quality assessments.

## Data Availability

The Landsat data was obtained from the USGS through the EarthExplorer (<https://earthexplorer.usgs.gov/>), and Sentinel-2 was downloaded from the ESA Sentinel Online (<https://sentinel.esa.int/web/sentinel/sentinel-data-access>). The rest of the data used to support the findings of this study are included within the article.

## Disclosure

The authors carried out the research while at the Department of Civil and Structural Engineering, Moi University, Kenya.

## Conflicts of Interest

The authors declare that they have no conflict of interest.

## References

- [1] N. Khatri and S. Tyagi, "Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas," *Frontiers in Life Science*, vol. 8, no. 1, pp. 23–39, 2014.
- [2] S. N. Kulshreshtha, "A global outlook for water resources to the year 2025," *Water Resources Management*, vol. 12, no. 3, pp. 167–184, 1998.
- [3] Y. O. Ouma, J. Waga, M. Okech, O. Lavisa, and D. Mbutia, "Estimation of Reservoir Bio-Optical Water Quality Parameters Using Smartphone Sensor Apps and Landsat ETM+: Review and Comparative Experimental Results," *Journal of Sensors*, vol. 2018, Article ID 3490757, 32 pages, 2018.
- [4] J. Boucher, K. C. Weathers, H. Norouzi, and B. Steele, "Assessing the effectiveness of Landsat 8 chlorophyll-*a* retrieval algorithms for regional freshwater monitoring," *Ecological Applications*, vol. 28, no. 4, pp. 1044–1054, 2018.
- [5] V. S. Martins, C. C. F. Barbosa, L. A. S. de Carvalho, D. S. F. Jorge, F. L. Lobo, and E. M. L. M. Novo, "Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to Amazon floodplain lakes," *Remote Sensing*, vol. 9, no. 4, p. 322, 2017.
- [6] T. Kutser, B. Paavel, C. Verpoorter et al., "Remote sensing of black lakes and Using 810 nm reflectance peak for retrieving water quality parameters of optically complex waters," *Remote Sensing*, vol. 8, no. 6, p. 497, 2016.
- [7] Y. O. Ouma, "Advancements in medium and high resolution Earth observation for land-surface imaging: evolutions, future trends and contributions to sustainable development," *Advances in Space Research*, vol. 57, no. 1, pp. 110–126, 2016.
- [8] M. Gholizadeh, A. Melesse, and L. Reddi, "A comprehensive review on water quality parameters estimation using remote sensing techniques," *Sensors*, vol. 16, no. 8, pp. 1298–1306, 2016.
- [9] Y. O. Ouma, C. O. Okuku, and E. N. Njau, "Use of artificial neural networks and multiple linear regression model for the prediction of dissolved oxygen in rivers: case study of hydrographic basin of River Nyando, Kenya," *Complexity*, vol. 2020, Article ID 9570789, 23 pages, 2020.
- [10] W. Zhu, Q. Yu, Y. Q. Tian, B. L. Becker, T. Zheng, and H. J. Carrick, "An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments," *Remote Sensing of Environment*, vol. 140, pp. 766–778, 2014.
- [11] F. Watanabe, E. Alcântara, T. Rodrigues, N. Imai, C. Barbosa, and L. Rotta, "Estimation of chlorophyll-*a* concentration and the trophic state of the Barra Bonita hydroelectric reservoir using OLI/Landsat-8 images," *International Journal of Environmental Research and Public Health*, vol. 12, no. 9, pp. 10391–10417, 2015.
- [12] M. Potes, G. Rodrigues, A. M. Penha et al., "Use of Sentinel 2 – MSI for water quality monitoring at Alqueva reservoir, Portugal," *Proceedings of the International Associations of Hydrological Sciences (IAHS)*, vol. 380, pp. 73–79, 2018.
- [13] K. Toming, T. Kutser, A. Laas, M. Sepp, B. Paavel, and T. Noges, "First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery," *Remote Sensing*, vol. 8, no. 8, p. 640, 2016.
- [14] W. He, S. Chen, X. Liu, and J. Chen, "Water quality monitoring in a slightly-polluted inland water body through remote sensing - case study of the Guanting Reservoir in Beijing, China," *Frontiers of Environmental Science & Engineering in China*, vol. 2, no. 2, pp. 163–171, 2008.
- [15] M. Potes, M. J. Costa, J. C. B. da Silva, A. M. Silva, and M. Morais, "Remote sensing of water quality parameters over Alqueva reservoir in the south of Portugal," *International Journal of Remote Sensing*, vol. 32, no. 12, pp. 3373–3388, 2011.
- [16] M. Potes, M. J. Costa, and R. Salgado, "Satellite remote sensing of water turbidity in Alqueva reservoir and implications on lake modelling," *Hydrology and Earth System Sciences*, vol. 16, no. 6, pp. 1623–1633, 2012.
- [17] IOCCG, *Remote sensing of ocean colour in coastal, and other optically-complex, waters*, S. Sathyendrath, Ed., Reports of the International Ocean-Colour Coordinating Group, No. 3, IOCCG, Dartmouth, Canada, 2000.
- [18] J. Li and D. Roy, "A global analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring," *Remote Sensing*, vol. 9, no. 9, p. 902, 2017.

- [19] A. N. Tyler, P. D. Hunter, E. Spyarakos, S. Groom, A. M. Constantinescu, and J. Kitchen, "Developments in earth observation for the assessment and monitoring of inland, transitional, coastal and shelf-sea waters," *Science of the Total Environment*, vol. 572, pp. 1307–1321, 2016.
- [20] P. Gernez, D. Doxaran, and L. Barillé, "Shellfish aquaculture from space: potential of Sentinel2 to monitor tide-driven changes in turbidity, chlorophyll concentration and oyster physiological response at the scale of an oyster farm," *Frontiers in Marine Science*, vol. 4, 2017.
- [21] H. Liu, Q. Li, T. Shi, S. Hu, G. Wu, and Q. Zhou, "Application of Sentinel 2 MSI images to retrieve suspended particulate matter concentrations in Poyang Lake," *Remote Sensing*, vol. 9, no. 7, p. 761, 2017.
- [22] K. Dörnhöfer, A. Göritz, P. Gege, B. Pflug, and N. Oppelt, "Water constituents and water depth retrieval from Sentinel-2A - a first evaluation in an oligotrophic lake," *Remote Sensing*, vol. 8, no. 11, p. 941, 2016.
- [23] U. N. T. Nguyen, L. T. H. Pham, and T. D. Dang, "An automatic water detection approach using Landsat 8 OLI and Google Earth Engine cloud computing to map lakes and reservoirs in New Zealand," *Environmental Monitoring and Assessment*, vol. 191, no. 4, 2019.
- [24] E. F. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sensing of Environment*, vol. 185, no. Iss 2, pp. 46–56, 2016.
- [25] H. K. Zhang, D. P. Roy, L. Yan et al., "Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences," *Remote Sensing of Environment*, vol. 215, pp. 482–494, 2018.
- [26] E. F. Vermote, D. Tanré, J. L. Deuze, M. Herman, and J. J. Morcette, "Second simulation of the satellite signal in the solar spectrum, 6S: an overview," *IEEE Geoscience and Remote Sensing Letters*, vol. 35, no. 3, pp. 675–686, 1997.
- [27] J. G. Masek, E. F. Vermote, N. E. Saleous et al., "A Landsat surface reflectance dataset for North America, 1990–2000," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 68–72, 2006.
- [28] D. P. Roy, J. Li, H. K. Zhang, L. Yan, H. Huang, and Z. Li, "Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance," *Remote Sensing of Environment*, vol. 199, pp. 25–38, 2017.
- [29] F. Gascon, C. Bouzinac, O. Thépaut et al., "Copernicus Sentinel-2A calibration and products validation status," *Remote Sensing*, vol. 9, no. 6, p. 584, 2017.
- [30] J. A. Barsi, B. Alhammoud, J. Czaplá-Myers et al., "Sentinel-2A MSI and Landsat-8 OLI radiometric cross comparison over desert sites," *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 822–837, 2018.
- [31] R. Chastain, I. Housman, J. Goldstein, M. Finco, and K. Tenneson, "Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States," *Remote Sensing of Environment*, vol. 221, pp. 274–285, 2019.
- [32] R. Shang and Z. Zhu, "Harmonizing Landsat 8 and Sentinel-2: a time-series-based reflectance adjustment approach," *Remote Sensing of Environment*, vol. 235, p. 111439, 2019.
- [33] M. Drusch, U. del Bello, S. Carlier et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sensing of Environment*, vol. 120, pp. 25–36, 2012.
- [34] J. R. Schott, *Remote Sensing the Image Chain Approach*, Oxford University Press, New York, 2nd Ed. edition, 2007.
- [35] B. A. Franz, S. W. Bailey, N. Kuring, and P. J. Werdell, "Ocean color measurements with the Operational Land Imager on Landsat-8: implementation and evaluation in SeaDAS," *Journal of Applied Remote Sensing*, vol. 9, no. 1, article 096070, 2015.
- [36] S. Yezpez, A. Laraque, J. M. Martinez et al., "Retrieval of suspended sediment concentrations using Landsat-8 OLI satellite images in the Orinoco River (Venezuela)," *Comptes Rendus Geoscience*, vol. 350, no. 1-2, pp. 20–30, 2018.
- [37] F. Mushtaq and M. G. Nee Lala, "Remote estimation of water quality parameters of Himalayan Lake (Kashmir) using Landsat 8 OLI imagery," *Geocarto International*, vol. 32, no. 3, pp. 274–285, 2016.
- [38] M. Bonansea, M. Ledesma, C. Rodriguez, and L. Pinotti, "Using new remote sensing satellites for assessing water quality in a reservoir," *Hydrological Sciences Journal*, vol. 64, no. 1, pp. 34–44, 2019.
- [39] F. Watanabe, E. Alcántara, T. Rodrigues, L. Rotta, N. Bernardo, and N. Imai, "Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/sentinel-2A (Barra Bonita reservoir, Brazil)," *Annals of the Academy of Brazilian Academy of Sciences*, vol. 90, 2 suppl 1, pp. 1987–2000, 2018.
- [40] V. Markogianni, D. Kalivas, G. Petropoulos, and E. Dimitriou, "Analysis on the feasibility of Landsat 8 imagery for water quality parameter assessment in an oligotrophic Mediterranean Lake," *International Journal of Geological and Environmental Engineering*, vol. 11, no. 9, pp. 906–914, 2017.
- [41] E. Kontopoulou, P. Kolokoussis, and K. Karantzalos, "Water quality estimation in Greek lakes from Landsat 8 multispectral satellite data," *European Water*, vol. 58, pp. 191–196, 2017.
- [42] L. C. González-Márquez, F. M. Torres-Bejarano, C. Rodríguez-Cuevas, A. C. Torregróza-Espinosa, and J. A. Sandoval-Romero, "Estimation of water quality parameters using Landsat 8 images: application to Playa Colorada Bay, Sinaloa, Mexico," *Applied Geomatics*, vol. 10, no. 2, pp. 147–158, 2018.
- [43] N. Lailia, F. Arafaha, L. Jaelania et al., "Development of water quality parameter retrieval algorithms for estimating total suspended solids and chlorophyll using Landsat 8 for Poteran Island waters," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-2/W2, 2015 Joint International Geoinformation Conference 2015, pp. 28–30, Kuala Lumpur, Malaysia, October 2015.
- [44] N. H. Quang, J. Sasaki, H. Higa, and N. H. Huan, "Spatiotemporal variation of turbidity based on Landsat 8 OLI in Cam Ranh Bay and Thuy Trieu Lagoon, Vietnam," *Water*, vol. 9, no. 8, p. 570, 2017.
- [45] L.-W. Liu and Y.-M. Wang, "Modelling reservoir turbidity using Landsat 8 satellite imagery by gene expression programming," *Water*, vol. 11, no. 7, p. 1479, 2019.
- [46] C. Roesler, J. Uitz, H. Claustre et al., "Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: a global analysis of WET labs ECO sensors," *Limnology and Oceanography: Methods*, vol. 15, no. 6, pp. 572–585, 2017.