Article

Estimation of Chlorophyll-*a* Concentration and Trophic States for an Inland Lake from Landsat 8 OLI Data: A Case of Nalban Lake of East Kolkata Wetland, India

Pulak Priti Patra ¹, Sourabh Kumar Dubey ^{1,*}, Raman Kumar Trivedi ¹, Sanjeev Kumar Sahu ² and Sangram Keshari Rout ¹

- ¹ Department of Aquatic Environment Management, Faculty of Fishery Sciences, West Bengal University of Animal and Fishery Sciences, Kolkata-700094, India; pulakpatra616@gmail.com (P.P.P.), trivediraman62@gmail.com (R.K.T.), skrout68@gmail.com (S.K.R.)
- ² ICAR-Central Inland Fisheries Research Institute, Barrackpore, Kolkata-700120, India; sksahu_2k@yahoo.com
- * Correspondence: sourabhkumardb@gmail.com; Tel.: +91-033-2432-8749

Abstract: 1) Landsat operational land imager (OLI) data and consequent laboratory measurements were used to predict Chlorophyll-a (Chl-a) concentration and the trophic states for an inland lake within the East Kolkata Wetland, India; 2) The most suitable band ratio was identified by performing Pearson correlation analysis between Chl-a concentrations and possible OLI band and band ratios from the study points; 3) The results showed highest correlation coefficient from the band ratio OLI5/OLI4 with an R value of 0.85. The prediction model was then developed by applying regression analysis between the band ratio OLI5/OLI4 and Chl-a concentration of the study points. The reflectance ratios of the validation points were given as input on the prediction model and the model output was considered as predicted Chl-a values of the validation points to check the efficiency of the prediction model. The regression model between laboratory-derived Chl-a value and model-fitted Chl-a value of the validation points revealed a high correlation with an R^2 value of 0.78. Trophic State Index (TSI) of the lake was also calculated from laboratory-derived Chl-a value and model-fitted Chl-a value of the validation points. The study presented a high correlation of TSI determined from predicted data with TSI from laboratory reference data (R = 0.88). The TSI values of the lake ranged from 65 to 75 which indicate that the lake is appeared to be eutrophic to hypereutrophic conditions. 4) This empirical study showed that Landsat 8 OLI imagery can be effectively applied to estimate Chl-a levels and trophic states for inland lakes.

Keywords: landsat 8 OLI; Nalban Lake; East Kolkata Wetland; chlorophyll-*a* prediction; study points; validation points

1. Introduction

Chlorophyll-*a* (chl-*a*) is a photosynthetic pigment that is found in all green floral components including algae [1]. An optimum concentration of phytoplankton and algae is crucial for a biologically productive and healthy lake as it is considered as the primary producer of aquatic ecosystem. Fish yield can be worth interpreted by primary production in lakes compared to any other preferred relationship between yield and environmental variables [2-3]. However, excessive concentration of Chlorophyll is undesirable as it inculcates eutrophic condition of the lake [4] and results in increment of phytoplankton standing crop [5]. The eutrophic condition also deteriorates water quality by external and internal nutrient loading which leads to disappearance of benthic fauna and adversely affects inhabitant fishes of aquatic body [6]. Potable water qualities, lives of

doi:10.20944/preprints201608.0149.v1

2 of 18

human and animal and recreational use of the lakes are adversely affected by excessive algal bloom posing a serious threat to society and environment [4, 7-12]. So, Chlorophyll estimation is regarded as an indicator of nutrient status of aquatic body and helpful for monitoring and managing the eutrophic lakes [13]. Throughout the world, incremental tendency of algal bloom in inland and coastal water body is a familiar phenomenon [4, 9, 12]. Therefore, in near future, implementation of monitoring plan for aquatic body (lake) may become a compulsion.

The Trophic State Index (TSI) is an important parameter to monitor the status of water quality as well as assessment of eutrophication [14]. Carlson [13] had developed the most suitable and acceptable methods for evaluating eutrophication in inland lakes. The index can be calculated from one of the several parameters such as Chl-*a*, Secchi disk depth, and total phosphorus. However, assessment methods of the eutrophic status of water bodies may be different as it largely depends on geographic locations, environmental issues, anthropogenic activities and the types of eutrophication [14]. For example, Shu [15] had developed a modified method of eutrophication assessment for the Chinese lake. In this method, chlorophyll-*a* was employed to calculate the trophic state index for lakes.

The traditional method of monitoring water quality is the manual collection of samples and laboratory analysis using in situ measurements. This process is not suitable to represent a synoptic spatio-temporal view of the large water bodies within a short time span and also not effective in terms of operational cost, manpower etc [16-17]. The remote sensing technique is capable to handle these problems and is useful in studying and monitoring the water quality parameters both at temporal and spatial scales without any field measurements [18]. The gradual development of optical and thermal sensors had made the remote sensing process an effective tool for extracting spatio-temporal information on water quality with high spectral and spatial resolution [19-20]. Satellite remote sensing is fast and relatively low operational cost process and can be used as a tool to derive spatio-temporal variability in lake water quality [21].

The 15 m and 30 m resolution of the Landsat 8 Operational Land Imager (OLI) combined with high global data availability, present a unique platform to provide the first and most up-to-date global inventory of the world's lakes and water quality information retrieval at high spatial resolution and positional accuracy using recent Landsat algorithms [22-24]. Miller et al. [25] had opined that the Landsat series provided approximately \$2.19 billion economic advantage per year that had been expensed over several fields of research. Brivio et al. [26] studied the relationship between Chl-a and remote sensing image data using Landsat TM images in Lake Garda (Italy). Duan et al. [14] studied estimation of Chl-a and TSI for inland lakes in northeast China from Landsat TM data. Giardino et al. [27] and Kim et al. [28] estimated Chl-a using Landsat 8 Operational Land Imager (OLI) in Lake Garda (Italy) and Fjord of Svalbard (Arctic sea) respectively. Lim and Choi [17] evaluated potential applications of the Landsat 8 OLI for estimating Chl-a in the Nakdong River (Korea). Theoretically, it is nearly impossible to detect and pin-point water quality data from remotely sensed images because of the water-atmosphere interface and unknown atmospheric aerosol [29]. So, empirical or semi-empirical approach should be introduced to quantify the water quality. Several researchers have successfully estimated different water quality parameters from the satellite band reflectance and developed empirical based predictive models for lake water quality [21, 30-33].

The objective of the present study is to determine the effectiveness of spectral reflectance values from the Landsat-8 OLI in respect to measurement of chl-*a* and Trophic State Index (TSI) in the Nalban Lake, East Kolkata Wetland (EKW), West Bengal, India. The study also optimizes the suitable OLI band or band combination for forecasting chl-*a* concentration and develops optimized regression models regionally.

2. Materials and Methods

2.1 Study Area

The Nalban Lake is situated inside the East Kolkata Wetland (EKW) (latitude 22°25' – 22°40' N, longitude 88°20' – 88°35' E), a complex of natural and man-made wetlands located at the eastern part of Kolkata, West Bengal, India (Figure 1). The EKW is one of the largest assemblages of sewage fed fish farms spread over an area of 12,500 ha. Because of its' immense ecological and socio-cultural importance, the EKW was designated as "Wetland of International Importance" under the Ramsar Convention on August 19, 2002. The water spread area of the Nalban Lake is about 126 ha. The detail physical characteristics of the lake were presented in Table 1. The lake sustains a diverse number of aquatic flora and fauna and also used for aquaculture purpose. Besides these, the lake is also used for tourism, recreation and sewage stabilization pond. This lake is now under the administrative control of State Fisheries Development Corporation Ltd, Government of West Bengal.

2.2 Water sampling procedure and measurements

The in-situ sampling was performed on the prefixed date when Landsat satellite overpasses the Kolkata region. The water sampling procedure was done in two phases and was restricted to surface water. In the first phase, eight sampling sites (S1-S8) were selected encompassing the entire lake surface area. These eight points were considered as 'study points' (Figure 1).

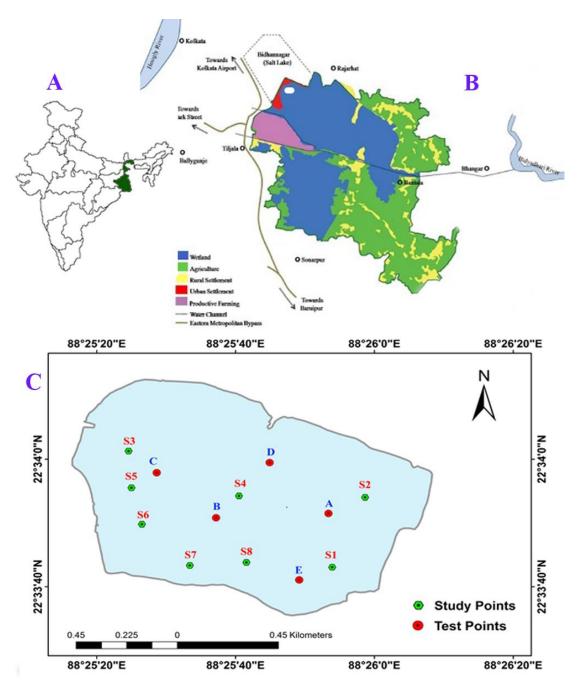


Figure 1. Map of the study area; A) India, B) East Kolkata Wetland, and C) Nalban Lake

Parameter	:	Values	
Elevation	:	1 m	
Lake Surface Area	:	126 ha	
Mean Depth	:	1.6 m	
Maximum Depth	:	2.5 m	
Maximum Width	:	0.87 km	
Maximum Length	:	1.47 km	
Residence Time	:	Perennial	

Table 1. Physical characteristics of Nalban Lake, India

Water samples were collected in amber bottles and were brought immediately to the laboratory for further analysis. The water sample was filtered using a glass micro-fibre filter (0.45 mm) followed by acetone extraction to estimate Chl-*a* which is a strong indicator for evaluating the biomass of algae. The concentration of Chl-*a* was determined by standard spectrophotometric method (HACH Spectrophotometer, DR 2800, Germany) after measuring the absorbance of the extracted dye at 663, 645, 630, and 750 nm. The Optical Density (OD) reading at 750 nm was a correction for turbidity and the actual amount of Chl-*a* was measured by the subtraction of the absorbance values at 750 nm from the others absorbance values. Finally, Chl-*a* concentration was calculated as per APHA [34] using the formula

Chl-a = 11.85(OD664) - 1.54(OD647) - 0.08(OD630)

The Chl-*a* concentration for the each study points (S1-S8) was tabulated in Table 2. The data obtained from these study points were considered as primary data for establishment of the prediction model. In the next phase, five 'validation points' were randomly selected from the lake surface (Figure 1) (Table 3). The sampling procedures were repeated for these five specific points on each satellite passing day and followed by same Chl-*a* estimation method. The sampling procedures were carried on 24th March 2015 to 27th May 2015. The further data sets were used for the validation of the prediction models. The co-ordinates of the both the study points and validation points were geo-located with the help of a hand held GPS recorder (Garmin eTrex® 10).

Table 2. Laboratory-measured Chlorophyll-a (Chl-a)) concentration of the study points in Nalban
Lake	

Study points	Sampling date	Latitude	Longitude	Chl-a (µg/L)
S1	16 th Nov 2014	22°33'43.06''N	88°25'53.92''E	34.94
S2	2 nd Dec 2014	22°33'54.02''N	88°25'58.66''E	40.04
S3	18 th Dec 2014	22°34'1.29''N	88°25'24.60''E	35.24
S4*	3 rd Jan 2015	22°33'54.25''N	88°25'40.50''E	24.30
S5*	19 th Jan 2015	22°33'55.52''N	88°25'25.01''E	22.24
S6	4 th Feb 2015	22°33'49.80''N	88°25'26.51''E	23.28
S7	20th Feb 2015	22°33'43.34"N	88°25'33.42''E	17.96
S 8	8th Mar2015	22°33'43.82''N	88°25'41.56''E	22.84
* These two stu	dy point data were	excluded due to >50%	6 cloud coverage in Lar	ndsat scenes

Table 3. The co-ordinates of the validation points over the Nalban Lake

Validation points	Latitude	Longitude	
Α	22°33'51.50''N	88°25'53.38''E	
В	22°33'50.82''N	88°25'37.19"E	
С	22°33'57.89"N	88°25'28.64''E	
D	22°33'59.48"N	88°25'44.90''E	
Ε	22°33'41.07''N	88°25'49.17''E	

2.3 Satellite Data

The cloud-free Landsat 8 OLI satellite images were used in the present study. It is an earth observation satellite that was developed by the National Aeronautics and Space Administration (NASA) and the Department of the Interior United States Geological Survey (USGS) in a partnership mode [35]. It consists of Operational Land Imager (OLI) sensor and Thermal Infrared Sensor (TRIS). Daily about 500 image scenes per day over the earth are captured and stored into the U.S. Landsat data archive at the USGS Earth Resource Observation and Science (EROS) Center, South Dakota [36]. The Landsat 8 OLI exhibits a higher resolution wavelength coverage than the Landsat 7 Enhanced Thematic Mapper plus (ETM+) bands due to the addition of a new coastal/aerosol band (430–450 nm) for detecting Chlorophyll and a new cirrus band (1.36–1.39 µm) for detecting clouds [37].

The satellite images used in this study were downloaded from the archive of USGS Landsat images (<u>http://earthexplorer.usgs.gov/</u>). This study utilized visible bands (blue, green, and red) and a near-infrared (NIR) band to determine correlations between Chl-*a* and spectral reflectance values. All image data from the Landsat 8 OLI were in GeoTIFF format provided by the US Geological Survey Earth Explorer. The details of Landsat bands used for model development were presented in Table 4. Scenes greater than 50% cloud coverage was excluded from model development. For the validation of the model, three Landsat scenes 24th March 2015, 27th April 2015 and 27th May 2015 were used.

Sensor	Resolution					Path/Row
	Bands	Wavelength	Spatial	Temporal	Radiometric	
		(nm)	(m)	(days)		
OLI	Band 2 (Blue)	450 – 515	30	16		138/44
	Band 3 (Green)	525 - 600	30	16	12 bit	
	Band 4 (Red)	630 - 680	30	16		
	Band 5 (NIR)	845 - 885	30	16		

Table 4. Detail description of Landsat 8 OLI satellite bands used in this study

2.4 Processing of satellite data

2.4.1 DN value extraction and conversion to TOA reflectance data

The digital number (DN) value was extracted from the satellite images for the co-ordinates of study points and validation points respectively with the help of TNT MIPS 2013 software (Version 15.0.0.533). In order to obtain top of the atmosphere (TOA) reflectance (ρ^{Λ}) recorded at the sensor, a conversion of the recorded signal is required. Therefore, the equation obtained from website (<u>http://landsat.usgs.gov/Landsat8_Using_Product.php</u>) was used to convert the DN value to TOA reflectance that is given below:

$$\varrho^{\lambda'} = M_{\varrho}^* Qcal + A_{\varrho} \tag{1}$$

Where, ρ^{λ} = TOA planetary reflectance, without correction for solar angle; M₀= Band specific multiplicative rescaling factor from metadata; A₀= Band specific additive rescaling factor from metadata; Qcal = Quantized and calibrated standard product pixel value

The TOA reflectance values were further rectified taking into account that water leaving radiance greatly varies depending on the solar angle. The value of local solar zenith angle and local sun elevation angle was taken from the metadata file provided with the Landsat data. The following equation obtained from the website (<u>http://landsat.usgs.gov/Landsat8_Using_Product.php</u>) was used for sun angle correction.

$$\varrho\lambda = \varrho\lambda'/\cos(\theta sz * \pi/180) = \varrho\lambda'/\sin(\theta se * \pi/180)$$
(2)

Where, $\rho \lambda$ = TOA planetary reflectance; θ se = Local sun elevation angle; θ sz = Local solar zenith angle.

2.4.2 Atmospheric Correction

Atmospheric correction is one most important step in deriving the land/water surface property from satellite data. The objective of atmospheric correction is to retrieve the surface reflectance (that characterized surface properties) from remotely sensed imagery by removing the atmospheric effects. The Darkest Pixel (DP) atmospheric correction method was applied to every satellite in the current study as it is very effective in the visible wavelength range [38]. This method assumes that any dark pixel on the scene that possesses the lowest DN value should have a zero reflectance [38]. Therefore its radiometric DN value represents the atmospheric additive effect [39-41]. So, the radiance value of that dark pixel needs to be subtracted from the radiance values of all pixels on that

band to remove the additive effect of atmospheric scattering. Here, the darkest point of an image was pointed out by analyzing the image histogram.

2.5 Development of prediction model

The statistical method is the most preferred approach to derive a correlation between spectral data and Chlorophyll concentration values [42-43]. Correlation between OLI brightness values and the water quality parameter by linear regression analysis was used in the present experiment [44]. Previous studies suggest that different band combinations (e.g. ratios, multiplication and average) which can be used to retrieve relationships with in-situ measurements [33, 45-46]. The Pearson correlation analysis was performed between possible bands/band combinations and the in situ Chl-*a* measurement of the study points to examine the relationship. Prediction models were then developed using the regression analysis.

2.6 Validation and accuracy assessment of model

To check the efficiency and prediction accuracy of the models, the validation process was also performed by field trial. The reflectance value of the validation points at the predictive band or band combination was given as input in the developed model and the model generated output values were considered as predictive water quality. Three statistical metrics was used to evaluate the efficiency of the predictive models: Bias, RMSE (Root Mean Square Error, in log space) and R^2 (regression, Type II) [47].

Bias is the measure of the difference between estimated value and the true value of the parameter being estimated. Its value ranges from 0 to ∞ . If the difference is zero, then it is called unbiased. The value close to zero indicates the better the performance of the model. It is calculated by the following formula [47]:

$$Bias = \frac{\sum_{i=1}^{n} (x_i^{estimated} - x_i^{measured})}{n}$$
(3)

The RMSE is a measure of the average magnitude of the error. Its value ranges from 0 to ∞ . Lower values of RMSE indicate better fit. RMSE is a good measure of how accurately the model predicts the response and is the most important criterion for fit if the main purpose of the model is a prediction. It was calculated by the following formula [47]:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left[\log(x_i^{estimated}) - \log(x_i^{measured})\right]^2}{n-2}}$$
(4)

Where $x_i^{measured}$ represents measured in situ water quality data at the validation sites and $x_i^{estimated}$ is the model estimated water quality and *n* is the number of validation points.

R-squared value (R^2) provides information about the goodness of fit of a model and helps to understand that the proximity of regression line to the real data points. It gives an idea how many data points lie within the results of the line formed by the regression equation [47]. Its value ranges from 0 to 1. Generally the higher the R^2 value, the better the model fits the data.

2.7 Estimation of Trophic State Index (TSI)

Trophic State Index (TSI) is a term which describes how "green" the lake is in respect of algal biomass content in the lake water. Carlson [13] developed a continuous scale 0 to 100 to express the trophic state of the lake based on either Secchi disk transparency, chl-*a* concentration or total phosphorus content. A TSI ranged from 40 to 50 can be assigned to the mesotrophic state, whereas values of more than 70 termed as hypereutrophic conditions [14-15, 48]. The following equation calculated the TSI based on Chl-*a* concentration that are given below [13]

TSI CHL=**10**
$$\left[6 - \frac{2.04 - 0.68 \ln(Chl - a)}{\ln 2}\right]$$
 (5)

2.7 Data Analysis

The details process and working flow of the study is given in Chart 1. The entire satellite data processing was done using the TNT MIPS 2013 (Ver 15.0.0.533) software. In order to establish the model, the statistical analysis was performed using IBM SPSS 20.0 statistical software [49]. The study area boundary map and spatial maps were prepared through ArcGIS software (10.2.2). The graphs and plots were generated using statistical software Origin® version 6.1 (OriginLab Corporation, Northampton, USA).

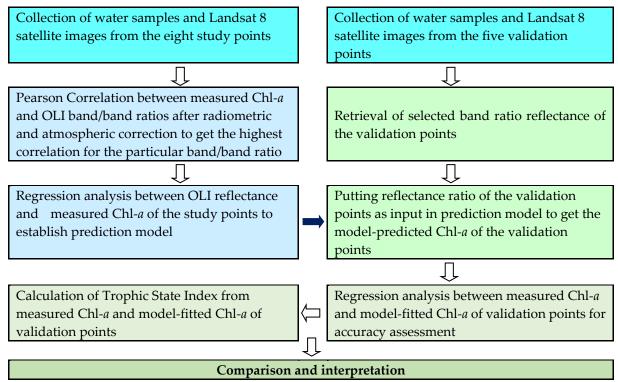


Chart 1. The detailed process and work flow of the study

3. Results and Discussion

In situ Chl-*a* concentrations of Nalban Lake, West Bengal, India varied during the entire study length and ranged from 17.96 μ g/L to 58.51 μ g/L. The highest value of Chl-*a* was observed in May 2015. Nalban Lake had Chl-*a* level of 33.58 - 35.94 μ g/L in March 2015 and 39.11 – 54.34 μ g/L in April 2015. Chl-*a* values of 50-250 μ g/L appear to be reasonable approximations of the range of phytoplankton biomass over which net primary production is maximized [5]. Boyd [50] found Chl-*a* values ranged 60-150 μ g/L as typical from productive fertilized fish and shrimp ponds. The relationship between Chl-*a* concentration and the trophic status of lakes and reservoirs is provided in Table 5.

Table 5. Relationship between Chlorophyll-*a* (μ g/L) concentrations and trophic condition of lakes and reservoirs (Adopted from Boyd [51])

Chlorophyll-a (µg/L	.)	
Annual mean	Annual maximum	Conditions
<2	<5	Oligotrophic, aesthetically pleasing, very low phytoplankton
		levels
2–5	5–15	Mesotrophic, some algal turbidity, reduced aesthetic appeal,
		oxygen depletion not likely
5–15	15–40	Mesotrophic, obvious algal turbidity, reduced aesthetic appeal,
		oxygen depletion likely
>15	>40	Eutrophic, high levels of phytoplankton growth, significantly
		reduced aesthetic appeal, serious oxygen depletion in bottom
		waters, reduction in other uses

3.1 Development of the prediction model

The Landsat visible bands ranging from blue (OLI 2) to green (OLI 4) and near-infrared band (NIR) (OLI 5) are commonly used for lake study and applied to obtain the relationship between the sub-surface reflectance and the bio-physical parameters (e.g. water transparency, Chl-*a* and total suspended solids etc.) of the water [52]. Pearson correlation analysis was performed to observe the relative strengths between Chl-*a* concentration and OLI bands (2-5) and their combinations of the study points. The Pearson correlation coefficients (*R*) between Chl-*a* concentration and various Landsat OLI bands and band ratios of the study points were depicted in Table 6. As shown in Table 6 the correlation between Chl-*a* concentration and OLI band ranged from 0.85 (B5/B4) to 0.49 (B5). The ratio between NIR band and red band (B5/B4) showed the highest correlation with Chl-*a* (*R*= 0.85; *p* <0.05). Lim et al. (2015) observed the correlation of OLI bands with Chl-*a* ranged from -0.64 (band 4) to -0.71 (band 5) and Chl-*a* concentration displayed a significant relationship (*R*= -0.71) with the NIR band at a significance level of *p* <0.01 Therefore, this band ratio (B5/B4) was considered as the best predictor of Chl-*a* concentration and further used for developing the prediction model.

Linear regression analysis was performed between the spectral reflectance values (B5/B4) and the in-situ measured Chl-*a* dataset of the study points to establish the predictive model. The analysis demonstrated a significant relationship ($R^2 = 0.72$) with the standard error of estimate (SEE) of 5.14 (µg/L) (Fig. 2). The regression equation is as follows

 $Chl-a_{C} = a + b (OL15/OL14) + c$

(6)

Where, *Chl-ac* is the concentration of Chl-*a*; (OLI5/OLI4) is the atmospherically corrected band ratio data; *a* and *b* are the regression coefficients and equals to -59.40 and 98.87 respectively.

Band Combinations	R	Band Combinations	R
B2	0.41	(B3+B5)/2	0.32
B3	0.16	(B2+B5)/2	0.45
B4	0.30	(B4+B3)/2	0.23
B5	0.49	(B4+B2)/2	0.37
B5/B2	-0.54	(B3+B2)/2	0.30
B5/B3	0.009	(B2+B3+B4)/3	0.30
B5/B4	0.85	(B2+B3+B5)/3	0.36
B4/B3	-0.26	(B3+B4+B5)/3	0.31
B4/B2	-0.64	(B2+B4+B5)/3	0.31
B3/B2	-0.71	(B2+B3+B4+B5)/4	0.35
(B4+B5)/2	0.40	The bold indicates highes	t correlation

Table 6. Pearson correlation coefficients (*R*) between Chlorophyll-*a* (Chl-*a*) concentration and various Landsat OLI bands and band ratios of the study points



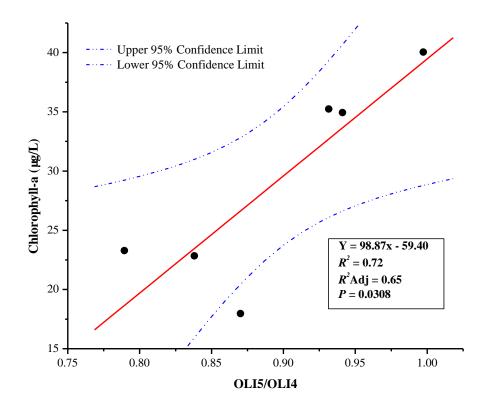


Figure 2. The regression line showing linear relationship between Chlorophyll-*a* (Chl-*a*) concentration and OLI5/OLI4 band ratio of the study points

Wang et al. [53] demonstrated that this kind of statistical analysis has been widely used to understand the relationship between in situ data and spectral reflectance values. Song et al. [54] used a correlation analysis to understand the relationship between water quality parameters and spectral reflectance values from the Landsat imagery. Chlorophyll absorbs energy in the wavelengths band at about blue and red, but strongly reflects energy in the wavelengths band of NIR [55]. But due to high pigment concentration and turbid condition of the inland water body, the water leaving radiance of the blue band is significantly small. So, the red band reflectance values are used instead of a blue band. Spectral bands near 676 nm (red band) have been widely used for the retrieval of Chl-a in shallow turbid coastal waters where the Chlorophyll concentration is comparatively high than open ocean [56-57]. The reflectance is high at infrared (IR) and NIR due to the combination of decreasing absorption by the Chlorophyll pigments and the increasing of absorption by water [58-59]. Gitelson et al. [60] observed that reflectance increases in the NIR beyond 700 nm due to increased scattering from algal biomass. Tebbs et al. [61] found that Landsat ETM + band ratio of NIR/red produced the best correlation with the in situ Chl-a measurement. Gitelson et al. [62] also successfully used NIR/red band ratio in turbid lake and reservoirs for predicting a high level of Chl-a concentration. Shen et al. [63] and Ruddick et al. [64] also successfully applied the spectral reflectance at Red-NIR region to detect and monitor harmful algal bloom at inland and coastal waters. Theologou et al. [65] also found a high correlation of Chl-a with red and NIR regions of Landsat 7 and 8 satellite for predicting water quality in inland shallow lakes.

3.2 Validation and accuracy assessment of prediction model

To predict the Chl-*a* using the Landsat 8 OLI data, the atmospherically corrected (OLI5/OLI4) band ratio of the validation points co-ordinates obtained from the Landsat scenes of March 2015, April 2015 and May 2015 were used. The reflectance ratios of the validation points were given as input to the prediction model equation (6) and the model output was considered as predicted Chl-*a* values of the validation points (satellite-derived). The regression model statistics between laboratory-derived Chl-*a* value and model-fitted Chl-*a* value of the validation points revealed a

strong correlation (R = 0.88; P < 0.0001) with a Bias of 10.04 µg/L; and an RMSE of 0.14 µg/L (Figure 3). The statistical interpretation showed that the scatter points were close to the 1:1 line without any outlier. Consistent with this study, Lim et al. (2015) estimated a stronger relationship between in situ Chl-*a* data and predicted values (R = 0.77). Giardino et al. [27] obtained 1.01 µg/L Chl-*a* through laboratory analysis and the average concentration through image data in Landsat 8 OLI showed 1.04 µg/L. Likewise, Lim and Choi [17] measured 17.50 µg/L, 21.10 µg/L and 16.00 µg/L concentration of Chl-*a* from Dalsung, Youngsan and Namji sites of Nakdong River, Korea and Landsat 8 OLI showed 17.55 µg/L, 15.95 µg/L, 20.98 µg/L concentration of Chl-*a* respectively. These are well consistent with the trend found in the present study emphasizing that the Landsat OLI is suitable to monitor the water quality in Nalban Lake.

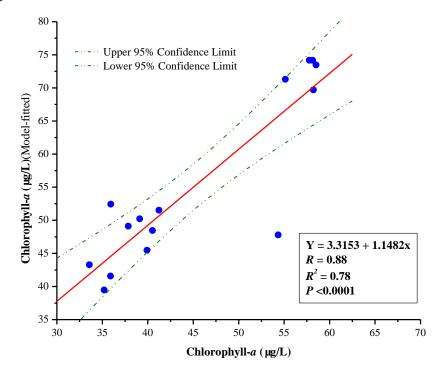


Figure 3.The regression between laboratory-derived Chlorophyll-*a* (Chl-*a*) concentration and model-fitted Chl-*a* value of the validation points

Figure 4 showed the spatial pattern of Chl-*a* concentration estimated from Landsat 8 OLI data from March to May 2015 in Nalban Lake. Chl-*a* concentration ranged from 39.47 to 74.17 μ g/L and the highest Chl-*a* concentration at the Nalban Lake was estimated on 27th May 2015. Based on the spatial distribution of Chl-*a*, the lake was likely in a state of eutrophic condition [51]. Although the values of Chl-*a* concentration were slightly overestimated using remote sensing techniques; this error may be attributed to a mismatch between the collection times of satellite based and ground-based measurements [17]. Therefore, the predicted value of Chl-*a* (satellite-derived) was further applied for the determination of TSI of the validation points.



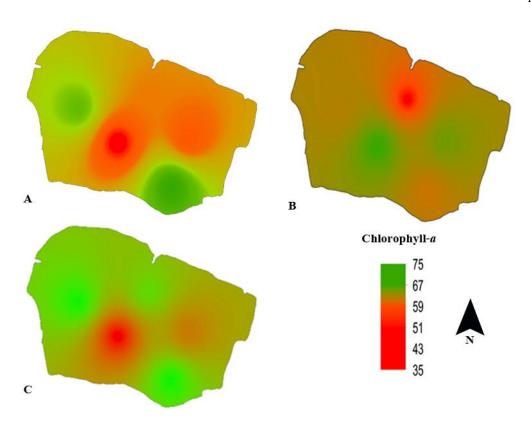


Figure 4. Spatial pattern of Chlorophyll-*a* (Chl-*a*) concentration estimated from Landsat 8 OLI data on A) March 2015, B) April 2015 and C) May 2015 in Nalban Lake

3.3 Trophic State Index (TSI) from field measurement and satellite data

Applying the equation (5), TSI of the water body was calculated from laboratory-derived Chl-a value and model-fitted Chl-a value of the validation points. Two methods give results of TSI for the consecutive three months at the close level and between 65 and 75. The comparison of these results (Figure 5) showed that the accuracy is satisfactory in this study which is very similar to the findings of other studies [14, 48]. In consistent with the present study, the TSI of the Lakes Bramin, Kagar and Schwarz from Germany was ranged from 50 to 70 [48]. Figure 6 gave a high correlation (R = 0.88; P<0.0001) of TSI determined from satellite-predicted data with TSI from laboratory reference data of the validation points with R^2 of 0.76 and RMSE of 1.11 mg/L in Nalban Lake. The spatial pattern of TSI estimated from Landsat 8 OLI data from March to May 2015 in Nalban Lake was depicted in Figure 7. TSI value ranged from 66.63 to 72.8 μ g/L and the highest TSI at the Nalban Lake was estimated on 27th May 2015. The TSI values indicated that the Nalban Lake is appeared to be eutrophic to hypereutrophic conditions [14-15, 48]. The Chl-a content of 7.25 µg/L results in an intermediate TSI of 50 and the standard deviation for Chl-a determination from the field spectra is about 7 µg/L [48]. The TSI is very sensitive for low Chl-a content thus the determination of TSI based on Chl-a content gives more reliable results for eutrophic than for mesotrophic lakes [48]. Therefore, the variation of TSI is well expressed and estimated through remote sensing technique especially from Landsat 8 OLI and the analysis of Chl-a using satellite remote sensing technique seems to be an appropriate method for the determination of trophic states for the wetlands and lakes.



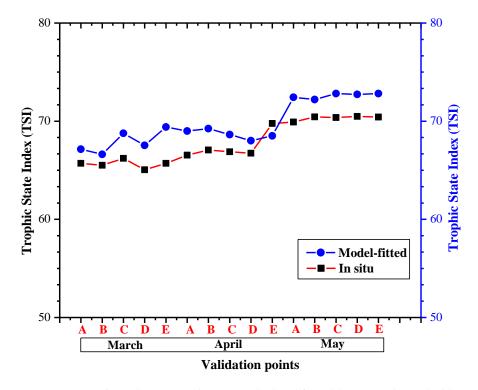


Figure 5. Comparisons of Trophic State Index (TSI) calculated from laboratory-derived Chlorophyll-*a* (Chl-*a*) and model-fitted Chl-*a* value of the validation points

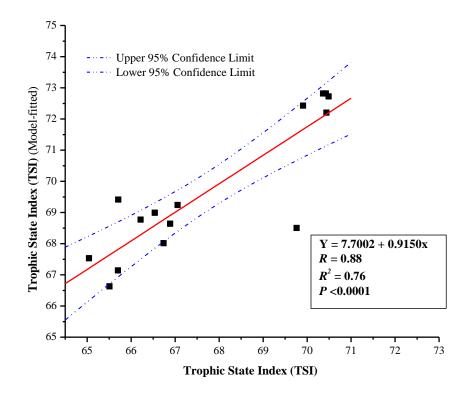


Figure 6 The regression between laboratory-derived Trohic State Index (TSI) values with model-fitted TSI values of the validation points



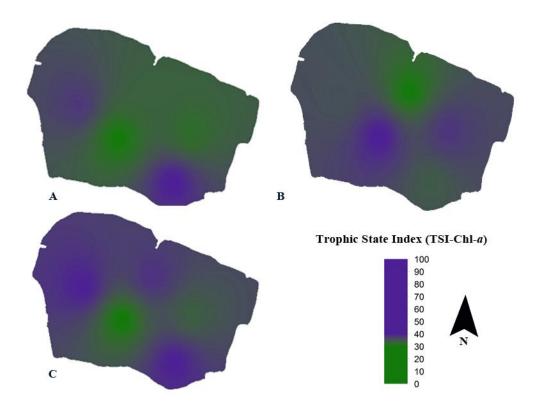


Figure 7. Spatial pattern of Trophic State Index (TSI) estimated from Landsat 8 OLI data on A) March 2015, B) April 2015 and C) May 2015 in Nalban Lake

5. Conclusions

The objective of the present study was to examine the potential application of the Landsat 8 satellite for estimating Chl-a concentration and evaluating the trophic states using remote sensing and statistical methods of the Nalban Lake situated within the Ramsar designated East Kolkata Wetland (EKW), India. As there was a small variation between satellite-estimated values and measured concentrations, this study demonstrates that Landsat 8 OLI data may provide a useful tool for investigating the spatio-temporal variability of Chl-a in surface waters. In the present study, NIR band and red band combinations (B5/B4) showed the strongest relationship with Chl-a and were also good predictors of the variance of Chl-a. As there was no significant difference in TSI measured from laboratory-derived Chl-a data and from the satellite-predicted Chl-a data, this research proves that Landsat data (especially red band and near-infrared band) is useful and reliable for investigating the trophic status of wetlands and lakes. The great advantage of using satellite imagery is the capability of their multi-temporal application covering a large area and low costs compared to laboratory analysis for the evaluation of the trophic states over several lakes. As the TSI value obtained from the present study depicted eutrophic to the hypereutrophic status of the Nalban Lake, this study suggests long-term monitoring of water quality conditions and recommends implementing suitable management plans to sustain the aquaculture and other activities over the lake. Further fine tuning of raw image data from the Landsat 8 OLI sensor through calibration and validation and further long-term research using remote sensing technologies will enable improvement of results in water quality of the Nalban Lake as well as East Kolkata Wetland, India.

Acknowledgments: The authors would like to acknowledge ICAR-Central Fisheries Research Institute, Barrackpore for helping in analyzing the satellite data. We are also thankful to Dr. Bijan Mandal (General Manager, SFDC) for giving permission to carry out research work in Nalban Lake. We are also grateful to Ms. Malancha Roy, School of Environmental Studies, Jadavpur University and Dr. Kaberi Samanta, School of Oceanographic Studies, Jadavpur University for improving the manuscript.

Author Contributions: P.P. and R.K.T. conceived and designed the experiments; P.P. performed the experiments; P.P. and S.K.S. analyzed the data; P.P. and S.K.D. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest

References

- 1. Wetzel, R.G. *Limnology: Lake and river ecosystems*, 3rd ed.; Academic Press: San Diego, California, 2001; pp. 1006.
- 2. Melack, J.H. Primary productivity and fish yields in tropical lakes. T. Am. Fish Soc. **1976**, 105, 575-580. http://dx.doi.org/10.1577/1548-8659(1976)105<575:PPAFYI>2.0.CO;2
- 3. Liang, Y.; Melack, J.M.; Wang, J. Primary production and fish yields in Chinese ponds and lakes. *T. Am. Fish Soc.* **1981**, 110, 346-350. http://dx.doi.org/10.1577/1548-8659(1981)110<346:PPAFYI>2.0.CO;2
- 4. Hallegraeff, G.M. Harmful algal blooms: A global review. In *Manual on harmful marine microalgae* Hallegraeff, G.M.; Anderson, D.M.; Cembella, A.D. Eds,; IOC Manuals and Guides No 33, UNESCO, Parris, France, 2003, pp 1-22.
- 5. Boyd, C.E.; Tucker, C.S. *Pond aquaculture water quality management*. Kluwer Academic publishers, London, 1998; pp. 700.
- Paerl, H.W. Nuisance phytoplankton blooms in coastal, estuarine and inland waters. *Limnol. Oceanogr.* 1988, 33, 823–843. http://dx.doi.org/10.4319/lo.1988.33.4part2.0823
- 7. Backer, L.C.; McGillicuddy, D.J. Jr. Harmful algal blooms. At the interface between coastal oceanography and human health. Oceanography **2006**, 19, 94-106. http://dx.doi.org/10.5670/oceanog.2006.72
- 8. Edler, L.; Fernö, S.; Lind, M.G.; Lundberg, R.; Nilsson, P.O. Mortality of dogs associated with a bloom of the cyanobacterium Nodularia spumigena in the Baltic Sea. *Ophelia* **1985**, 24, 103-109. http://dx.doi.org/10.1080/00785236.1985.10426623
- 9. Glibert, P.M.; Seitzinger, S.; Heil, C.A.; Burkholder, J.M.; Parrow, M.W.; Codispoti, L.A.; Kelly, V. The role of eutrophication in the global proliferation of harmful algal blooms: New perspectives and new approaches. *Oceanography* **2005**, 18,198-209.
- 10. Horner, R.A.; Garrison, D.L.; Plumley, F.G. Harmful algal blooms and red tide problems on the US west coast. *Limnol. Oceanogr.* **1997**, 42, 1076-1088. http://dx.doi.org/10.4319/lo.1997.42.5_part_2.1076
- 11. Landsberg, J.H. The effects of harmful algal blooms on aquatic organisms. *Rev. Fish. Sci.* **2002**, 10, 113-390. http://dx.doi.org/10.1080/20026491051695
- Sellner, K.G.; Douchette, G.J.; Kirkpatrick, G.J. Harmful algal blooms: causes, impacts and detection. J. Ind. Microbiol. Biotechnol. 2003, 30, 383-406. http://dx.doi.org/10.1007/s10295-003-0074-9
- 13. Carlson, R.E. A trophic state index for lakes. *Limnol. Oceanogr.* **1977**, **22**, 361-369. http://dx.doi.org/10.4319/lo.1977.22.2.0361
- Duan, H.; Zhang, Y.; Zhang, B.; Song, K.; Wang, Z.; Liu, D. Estimation of chlorophyll-a concentration and trophic states for inland lakes in Northeast China from Landsat TM data and field spectral measurements. *Int. J. Remote Sens.* 2008, 29,767-786. http://dx.doi.org/10.1080/01431160701355249
- 15. Shu, J.H. Assessment of eutrophication in main lakes of China. Oceanol. Limnol. Sinica. 1993, 24, 616–620.
- 16. Giardino, C.; Pepe, M.; Brivio, P.A.; Ghezzi, P.; Zilioli, E. Detecting chlorophyll, Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *Sci. Total Environ.* **2001**, 268, 19–29. http://dx.doi.org/10.1016/S0048-9697(00)00692-6
- 17. Lim, J.; Choi, M. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environ. Monit. Assess.* **2015**, 187, 384. http://dx.doi.org/10.1007/s10661-015-4616-1
- 18. Senay, G.B.; Shafique, N.A.; Autrey, B.C.; Fulk, F.; Cormier, S.M. The selection of narrow wavebands for optimizing water quality monitoring on the Great Miami River, Ohio using Hyperspectral remote Sensor data. *J Spat. Hydrol.* **2001**, *1*, 1-22.

- Nas, B.; Ekercin, S.; Karabörk, H.; Berktay, A.; Mulla, D.J. An application of Landsat-5TM image data for water quality mapping in Lake Beysehir, Turkey. *Water Air Soil Pollut*. 2010, 212, 183–197. http://dx.doi.org/10.1007/s11270-010-0331-2
- 20. Ritchie, J.C.; Zimba, P.V.; Everitt, J.H. Remote sensing techniques to assess water quality. *Photogramm. Eng. Rem. S.* **2003**, 69, 695–704.
- 21. Zilioli, E.; Brivio, P.A. The satellite derived optical information for the comparative assessment of lacustrine water quality. *Sci. Total Environ.* **1997**, 196, 229–245. http://dx.doi.org/10.1016/S0048-9697(96)05411-3
- 22. Li, J.; Sheng, Y. An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas. *Int. J. Remote Sens.* **2012**, 33, 5194–5213. http://dx.doi.org/10.1080/01431161.2012.657370
- 23. Sheng, Y.; Li, J. Satellite-observed endorheic lake dynamics across the Tibetan Plateau between circa 1976 and 2000. In: Remote sensing of protected lands; Wang, Y.Q., Ed, CRC Press: New York, USA, 2011, pp. 305–319.
- 24. Smith, L.C.; Sheng, Y.; MacDonald, G.M.; Hinzman, L.D. Disappearing Arctic lakes. *Science* 2005, 308, 1429. http://dx.doi.org/10.1126/science.1108142
- Miller, H.M.; Sexton, N.R.; Koontz, L.; Loomis, J.; Koontz, S.R.; Hermans, C. 2011 The users, uses, and value of Landsat and other moderate-resolution satellite imagery in the United States Executive report: U.S. Geological Survey Open-File Report2011-1031, 43p. Available online: https://pubs.usgs.gov/of/2011/1031/pdf/OF11-1031.pdf (accessed on 13 August 2016).
- 26. Brivio, P.; Giardino, C.; Zilioli, E. Determination of chlorophyll concentration changes in Lake Garda using an image-based radiative transfer code for Landsat TM images. *Int. J. Remote Sens.* **2001**, 22, 487–502.
- 27. Giardino, C.; Bresciani, M.; Cazzanigq, I.; Schenk, K.; Rieger, P.; Braga, F.; Matta, E.; Brando, V. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of Lake Garda. *Sensors* **2014**, 14, 24116–24131.
- 28. Kim, S.I.; Kim, H.C.; Hyun, C.U. High Resolution Ocean Color products estimation in Fjord of Svalbard, Arctic sea using Landsat 8 OLI. *Korean J. Remote Sens.* **2014**, 30, 809–816.
- 29. Sturm, B. The atmospheric correction of remotely sensed data and the quantitative determination of suspended matter in marine water surface layers. In *Remote sensing in meteorology, oceanography, and hydrology;* Cracknell, A.P. Ed.; Chichester: Ellis Horwood Ltd, UK, 1981, pp. 163–197.
- 30. Dekker, A.; Peters, S. The use of the Thematic Mapper for the analysis of eutrophic lakes: A case study in the Netherlands. *Int. J. Remote Sens.* **1993**, 14, 799-821. http://dx.doi.org/10.1080/01431169308904379
- Giardino, C.; Brando, V.E.; Dekker, A.G.; Strombeck, N.; Candiani, G. Assessment of water quality in Lake Garda (Italy) using Hyperion. *Remote Sens. Environ.* 2007, 109, 183-195. http://dx.doi.org/ 10.1016/j.rse.2006.12.017
- 32. Kallio, K. Remote sensing as a tool for monitoring lake water quality In *Water quality measurements series: Hydrological and limnological aspects of lake monitoring*; Heinonen, P.; Ziglio, G.; Van Der Beken, A. Eds.; John Wiley and Sons Ltd., Chichester, UK, 2000, pp. 237-245.
- 33. Lathrop, R. Landsat thematic mapper monitoring of turbid inland water quality. *Photogramm. Eng. Remote Sens.* **1992**, 58, 465-470.
- 34. *Standard methods for the examination of water and wastewater*. 22nd ed.; Rice, E.W.; Baird, R.B.; Eton, A.D.; Clesceri, L.S. Eds,; American Public Health Association, American Water Works Association, and Water Environment Federation, Washington, DC, 2012.
- 35. Irons, J.R.; Loveland , T.R. Eighth Landsat satellite becomes operational. *Photogramm. Eng. Remote Sens.* **2013**, 79, 398–401.
- Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.B.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E.; Nemani, R.; Oreopoulos, L.; Schott, J.; Thenkabail, P.S.; Vermote, E.F.; Vogelmann, J.; Wulder, M.A.; Wynne, R. Free access to Landsat imagery. *Science* 2008, 320, 1011. http://dx.doi.org/10.1126/science.320.5879.1011a

- Roy, D.P.; Qin, Y.; Kovalskyy, V.; Vermote, E.F.; Ju. J; Egorov, A.; Hansen, M.C.; Kommareddy, I.; Yan, L. Conterminous United States demonstration and characterization of MODIS-based Landsat ETM+ atmospheric correction. *Remote Sens. Environ.* 2014, 140, 433–449. http://dx.doi.org/10.1016/j.rse.2014.02.001
- Hadjimitsis, D.G.; Clayton, C.R.I.; Hope, V.S. An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs. *Int. J. Remote Sens.* 2004, 25, 3651–3674. http://dx.doi.org/10.1080/01431160310001647993
- 39. Campbell, J.B. Introduction to remote sensing, 3rd ed.; CRC Press, London, 2002; pp. 654.
- 40. Crane, R.B. Pre-processing techniques to reduce atmospheric and sensor variability in multi-spectral scanner data. In Proceedings of the 7th International Symposium on Remote Sensing of Environment, Ann Arbor, University of Michigan, USA, 1971.
- 41. Crippen, R.E. The regression intersection method of adjusting image data for band rationing. *Int. J. Remote Sens.* **1987**, *8*, 137–155. http://dx.doi.org/10.1080/01431168708948622
- 42. Baban, S.M.J. Detecting Water-quality parameters in the Norfolk Broads, UK, using Landsat Imagery. *Int. J. Remote Sens.* **1993**, 14, 1247-1267. http://dx.doi.org/10.1080/01431169308953955
- Mayo, M.; Gitelson, A.; Yacobi, Y.Z.; Ben-Avraham, Z. Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. *Int. J. Remote Sens.* 1995, 16, 175-182. http://dx.doi.org/10.1080/01431169508954386
- 44. Bilge, F.; Dogeroglu, T.; Ayday, C. (1997) Mapping of water quality parameters by using Landsat images in Porsuk Dam Lake, Eskisehir, Turkey. In: Proceedings of the 11 th International Symposium on Geology and Environment, Chamber of Geological Engineers of UCEAT, Turkey, Yilmazer, I. Ed.; pp. 101-107.
- Lavery, P.; Pattiaratchi, C.; Wyllie, A.; Hick, P. Water quality monitoring in estuarine waters using the landsat thematic mapper. *Remote Sens. Environ.* 1993, 46, 268-280. http://dx.doi.org/10.1016/0034-4257(93)90047-2
- 46. Kloiber, S.M.; Brezonik, P.L.; Olmanson, L.G.; Bauer, M.E. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sens. Environ.* **2002**, 82, 38-47.
- 47. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. In: Lee, Z.P. Ed.; Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada, 2006. Available online: <u>http://www.ioccg.org/reports/report5.pdf</u> (accessed on 14 August 2016).
- Thiemann, S.; Kaufmann, H. Determination of Chlorophyll Content and Trophic State of Lakes Using Field Spectrometer and IRS-1C Satellite Data in the Mecklenburg Lake District, Germany. *Remote Sens. Environ.* 2000, 73, 227–235. http://dx.doi.org/10.1016/S0034-4257(00)00097-3
- 49. IBM Corp (2011) IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp
- 50. Boyd, C.E. *Water Quality in ponds for aquaculture*. Auburn, A.L: Auburn University Alabama Agricultural Experiment Station, 1990.
- 51. Boyd, C.E. *Water Quality: An Introduction*. 2nd ed.; Springer International Publishing, AG Switzerland, 2015, pp. 357.
- 52. Sass, G.Z.; Creed, I.F.; Bayley, S.E.; Devito, K.J. Understanding variation in trophic status of lakes on the Boreal Plain: A 20 year retrospective using Landsat TM imagery. *Remote Sens. Environ.* **2007**, 109, 127–141. http://dx.doi.org/10.1016/j.rse.2006.12.010
- 53. Wang, F.; Han, L.; Kung, H.T.; Van Arsdale, R. Applications of Landsat5 TM imagery in assessing and mapping water quality in Reelfoot Lake, Tennessee. *Int. J. Remote Sens.* **2006**, 27, 5269–5283. http://dx.doi.org/10.1080/01431160500191704
- Song, K.; Wang, Z.; Blackwell, J.; Zhang, B.; Li, F.;, Zhang, Y.; Jiang, G. Water quality monitoring using Landsat Themate Mapper data with empirical algorithms in Chagan Lake, China. *J. Appl. Remote Sens.* 2011, 5: 053506–053516. http://dx.doi.org/10.1117/1.3559497
- 55. Lillesand, M.T.; Kiefer, R.W.; Chipman, J.W. Remote sensing and image interpretation. 7th ed.; Wiley, Hoboken, New Jersey, U.S.A., 2015, pp. 736.

- 56. Gurlin, D.; Gitelson, A.A.; Moses, W.J. Remote estimation of chl-a concentration in turbid productive waters: return to a simple two-band NIR-red model? *Remote Sens. Environ.* **2011**, 115, 3479–3490. http://dx.doi.org/10.1016/j.rse.2011.08.011
- 57. Odermatt, D.; Gietlson, A.; Brando, V.E.; Schaepman, M. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sens. Environ.* **2012**, 118, 116–126. http://dx.doi.org/10.1016/j.rse.2011.11.013
- Sellner, K.G. Physiology, ecology, and toxic properties of marine cyanobacteria blooms. *Limnol. Oceanogr.* 1997, 42, 1089–1104.
- 59. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127-150. http://dx.doi.org/10.1016/0034-4257(79)90013-0
- 60. Gitelson, A.; Schalles, J.F.; Rundquist, D.C.; Schiebie, F.R.; Yacobi, Y.Z. Comparative reflectance properties of algal cultures with manipulated densities. *J. Appl. Phycol.* **1999**, 11, 345-354. http://dx.doi.org/10.1023/A:1008143902418
- 61. Tebbs, E.J.; Remedios, J.J.; Harpar, D.M. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypereutrophic, saline-alkaline, flamingo lake, using Landsat ETM. *Remote Sens. Environ.* **2013**, 135, 92-106. http://dx.doi.org/10.1016/j.rse.2013.03.024
- Gitelson, A.A.; Dall'Olmo, G.; Moses, W.M.; Rundquist, D.C.; Barrow, T.; Fisher, T.R.; Gurlin, D.; Holz, H. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: validation. *Remote Sens. Environ.* 2008, 112, 3582-3593. http://dx.doi.org/10.1016/j.rse.2008.04.015
- 63. Shen, L.; Xu, H.; Guo, X. Satellite remote sensing of harmful algal blooms (HAB) and a potential synthesized framework. Sensors **2012**, 12, 7778–7803.
- 64. Ruddick, K.; Lacroix, G.; Park, Y.; Rousseau, V.; De Cauwer, V.; Sterckx, S. Overview of ocean colour: theoretical background, sensors and applicability to detection and monitoring of harmful algal blooms (capabilities and limitations). In: *Real-Time coastal observing systems for marine ecosystem dynamics and harmful algal blooms: Theory, instrumentation and modeling*, Babin, M.; Roesler, C.S.; Cullen, J.J. Eds.; UNESCO, Paris, France, 2008, pp. 331–383.
- 65. Theologou, I.; Patelaki, M.; Karantzalos, K. (2015) Can single empirical algorithms accurately predict inland Shallow water quality status from high resolution, multi-sensor, Multi-temporal satellite data? Proceedings of The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7/W3, 36th International Symposium on Remote Sensing of Environment, 11–15 May 2015, Berlin, Germany.



© 2016 by the authors; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).