Chlorophyll-a Estimation in Turbid Waters Using Combined SAR Data With Hyperspectral Reflectance Data: A Case Study in Lake Taihu, China

Yuanzhi Zhang, *Member, IEEE*, Martti Hallikainen, *Fellow, IEEE*, Hongsheng Zhang, *Member, IEEE*, Hongtao Duan, Yu Li, *Member, IEEE*, and X. San Liang

***Abstract*—The estimation of chlorophyll-a (chl-a) concentration remains a great challenge in turbid waters due to their complex optical conditions. To improve chl-a estimation, this study aims to determine whether combined use of polarimetric synthetic-aperture radar (SAR) data has potential for improving the chl-a estimation from hyperspectral sensing reflectance for turbid waters such as those found in Lake Taihu, China. *In situ* measurements of hyperspectral reflectance data and water samples were collected over the lake corresponding to ENVISAT ASAR data. Semiempirical (two-band and three-band models) and empirical [multiple linear regression (MLR) and multilayer perceptron network (MLP)] models are compared to estimate the chl-a concentration from *in situ* hyperspectral reflectance and SAR data. The results show that there is a general underestimation of chl-a for concentrations higher than 26 ug/L, which is probably caused by the large spatial variation of chl-a in the study area. The results also demonstrate that the MLR model performs in a more stable manner than the MLP network does, while MLP underestimates low and high areas of chl-a concentrations in the lake. On the other hand, due to the availability of one scenic SAR data on the same day, our results show that the additional use of SAR data improved chl-a estimation very slightly in this case study, although the performance of vertical/vertical polarization SAR data was better than that of horizontal/horizontal polarization data in chl-a estimation. Potential future work in this subject could explore other measures of mutual information between SAR and hyperspectral optical data beyond the correlation and regression techniques described. Therefore, it is still necessary to apply more SAR data in varied turbid waters in the near future to determine how SAR data can be useful in the improvement of chl-a estimation.**

***Index Terms*—Chlorophyll-a estimation, hyperspectral reflectance data, polarimetric SAR, turbid waters.**

Manuscript received May 26, 2017; revised August 28, 2017; accepted December 16, 2017. Date of publication February 7, 2018; date of current version April 11, 2018. This work was supported in part by the "National Key Research and Development Program of China (2016YFB0501501)", in part by the "2015 Innovation Program for Research and Entrepreneurship Teams, Jiangsu Province, China", in part by the "Provincial Natural Science Foundation of Jiangsu of China (BK20160049)," and in part by the "Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD)". *(Corresponding author: Yuanzhi Zhang.)*

Y. Zhang is with Nanjing University of Information Science and Technology, School of Marine Sciences, Nanjing 210044, China, and also with the Chinese University of Hong Kong, Shenzhen Research Institute, Shenzhen 518057, China (e-mail: yuanzhizhang@cuhk.edu.hk).

X. San Liang is with Nanjing University of Information Science and Technology, School of Marine Sciences, Nanjing 210044, China (e-mail: sanliang@courant.nyu.edu).

M. Hallikainen is with the Aalto University, Department of Radio Science and Engineering, School of Electrical Engineering, Aalto 00076, Finland (e-mail: martti.hallikainen@aalto.fi).

H. Zhang is with the Chinese University of Hong Kong, Shenzhen Research Institute, Shenzhen 518057, China (e-mail: stevenzhang@cuhk.edu.hk).

H. Duan is with the Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China (e-mail: htduan@niglas.ac.cn).

Y. Li is with the Beijing University of Technology, Faculty of Information Technology, Beijing 100124, China (e-mail: yuli@bjut.edu.cn)

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2017.2789247

I. Introduction

CYANOBACTERIAL harmful algal blooms have increased in both fresh and marine waters throughout the world [1], including such important waters as Lake Erie (North America), Lake Victoria (Africa), Lake Loosdrecht (Europe). and Lake Taihu (Asia) [2]–[8]. Many genera of cyanobacteria are known to produce a wide variety of toxins with negative effects on human health and aquatic life [9]; overall, algal blooms pose a major threat to drinking and irrigation water supplies, fishing, and recreational use of surface waters worldwide [10].

Optical satellite remote sensing has been used for bloom assessment in many freshwater lakes [11]–[16]. A widely used proxy for bloom monitoring is the concentration of chlorophyll-a (chl-a) [17]–[20]. In fact, researchers have proposed a number of approaches to quantitatively estimate the concentration of chl-a using remotely sensed data [21]. These methods can be grouped into three categories: a bio-optical approach based on modeling the inherent optical properties (IOPs) and apparent optical properties of water constituents [22], an empirical approach such as using various regression methods, and a semiempirical approach such as the three-band model [23], [24]. Bio-optical models are considered to be more stable and universal since they have physical interpretations; however, they are hindered by IOPs that can be difficult to measure in water quality studies. As a result, semiempirical and empirical models are still widely used to estimate chl-a concentration [25]. Semiempirical models are designed as various band ratio algorithms or as a three-band model of the optical reflectance at different wavelengths mainly in the near-infrared (NIR)–red bands, such as around 660–670 nm, 700–710 nm, and 730–750 nm [26]–[32]. Empirical models such as the artificial neural network, the genetic algorithm, and the support vector machine (SVM) are also employed to estimate pigment concentration [33]–[36].

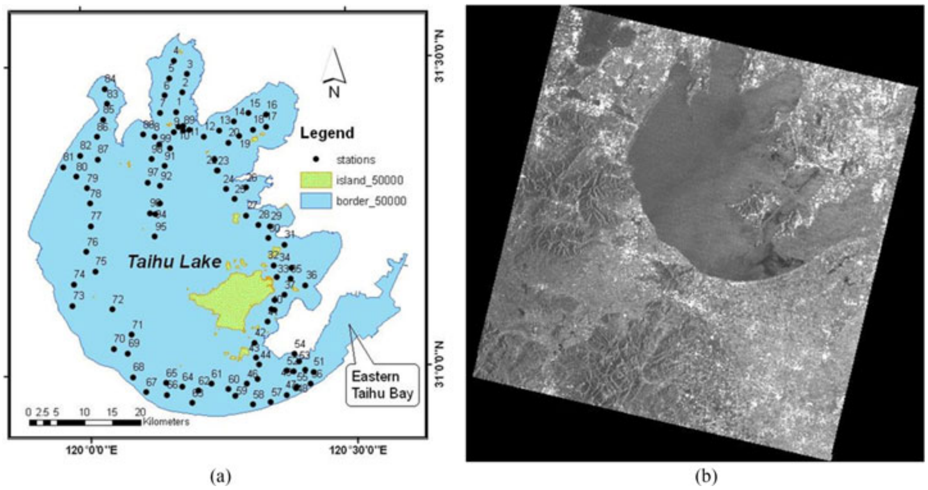

Fig. 1. Locations of sampling stations in Lake Taihu for *in situ* data and ASAR data. (a) Locations of stations. (b) ASAR data.

However, retrieving chl-a concentration in turbid lakes still poses a challenge due to the complex optical constituents of these waters [37], [38]. Moreover, uncertain atmospheric conditions (e.g., cloud cover) seriously affect the usefulness of optical satellite data in monitoring surface water quality [12]. In contrast, synthetic-aperture radar (SAR), working on the microwave band, is much less influenced by weather conditions and has proven capable of working all day and in all weather conditions [36], [39]. Microwaves are sensitive to the geometric configuration of the earth's surface, such as the roughness of land and water surface [40], [41]. In addition, microwaves are sensitive to the electromagnetic conductivity of the water surface [42]. This means that high chl-a concentrations influence the electromagnetic conductivity of the water surface since they change the components of the water surface [36] in the case of high chl-a concentrations. Moreover, when the water surface makes contact with wind, different chl-a concentrations and wind speeds and directions would lead to varying roughness of the water surface [43]. Therefore, SAR data holds potential for chl-a retrieval because of unique working characteristics: 1) SAR images mainly reflect the surface's geometric features (e.g., surface roughness), which relate to algal blooms since large quantities of algae change the roughness of the water surface; 2) Radar backscattering signals carry information about surface materials that change the permittivity (dielectric constant) [44], [45] of the surface water, and these materials potentially include pigments of phytoplankton or other water pollutants [36]. Nevertheless, researchers have paid little attention to the use of SAR data for chl-a estimation [43], and past studies regarding a SAR data-based methodology is insufficient [46], [47]. So far, there has been little comparison between SAR-based chl-a estimation methods and conventional models using optical remote-sensing data in turbid waters such as Lake Taihu.

The objective of this study is to investigate the potential of SAR data as a complementary source for hyperspectral remote-sensing data to gauge chl-a concentration and compare SAR data with information from conventional semianalytical chl-a retrieval models. The performance of two polarizations, vertical transmitting/vertical receiving (VV) and horizontal transmitting/horizontal receiving (HH), were examined in the study area of Lake Taihu in China, as it is well known for its increasing algal blooms and related increasing threats to millions of people. To estimate chl-a concentration, two groups of methods were employed. One group involved semiempirical models (i.e., two-band and three-band models), in which only optical data were applied to estimate chl-a. The other group comprised empirical methods [e.g., multiple linear regression (MLR) and multilayer perceptron network (MLP)]. The empirical methods are used to estimate chl-a concentration in two ways: solely optical data versus both optical and SAR data.

II. Data and Methods

A. Study Area

Lake Taihu (31.15 N, 120.15 E) is the third largest freshwater lake in China, with an area of 2338 km^2 and a mean depth of 1.9 meters [48]. In recent years, Lake Taihu has experienced significant pollution due to rapid economic growth in the surrounding region. Increasing eutrophication and recurring algal blooms pose a significant threat to the millions of people who rely on the lakes for their drinking water supply [49], [50]. A number of efforts have been put toward algal bloom study in Lake Taihu [51], [52]. In previous studies, various methods were examined and tested, including semianalytical models [51] such as three-band and four-band models, and empirical methods [52] such as the SVM. However, accurate monitoring of the algal blooms in Lake Taihu remains a challenging task due to the complex water conditions in this area [16], [17]. Therefore, combined use of optical and SAR data for chl-a concentration retrieval should be meaningful in monitoring algal blooms in Lake Taihu.

B. Field Measurements and Satellite Data

In this study, hyperspectral optical data and polarimetric SAR data were collected to assess the potential of SAR data in the estimation of chl-a. First, water samples and optical data were collected and selected on October 17, 2008 (see Fig. 1). The

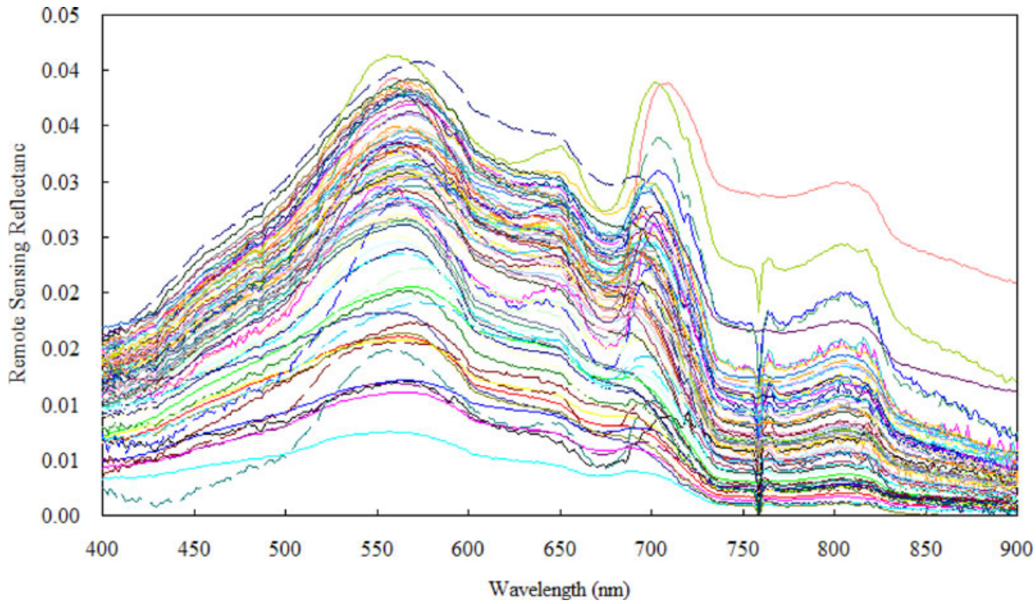

Fig. 2. Measured hyperspectral reflectance of water samples from 75 stations.

average wind speed was 6.53 m/s, with wind direction of 73.66° (northeast wind). The temperature was 11.73 °C, and the relative humidity was 67.88%.

At each station, hyperspectral remote-sensing reflectance ($R_{\text{rs}}(\lambda)$) was measured with a hand-held spectrometer, following NASA's Ocean Optics protocols [53] to calibrate and process the spectral measurements. Water samples were collected at the surface (to a depth of about 30 cm) with a standard 2-liter polyethylene water-fetching instrument immediately after R_{rs} measurement. The samples were stored with ice bags before conducting pigment absorption and concentration measurements in the laboratory. Chl-a was measured spectrophotometrically. Absorbance at 665 and 750 nm of the extract was measured with a UV-2401 spectrophotometer (Shimadzu Corp., Japan), and chl-a was calculated against filtered water as a reference [17]. In general, chl-a concentration in Lake Taihu varies at 3–108 ug/L seasonally, reaching high levels in summer and autumn, but was found at low levels in winter and early spring [2].

One scenic ASAR image in the alternating polarization (AP) mode (with VV and HH) was collected on October 17, 2008. The AP mode works with an incident angle at about 23.24°. The frequency was 5.331 GHz (C-Band), with a spatial resolution of 12.5 m. The two pieces of dual-polarization ASAR data were obtained with the same incident angle (only one look). The SAR data were geocoded, and water sampling stations were mapped in the SAR image according to their coordinates using handed GPS measurements [54]. Finally, 75 water sampling data and corresponding polarimetric SAR data were used for chl-a estimation after 25-station data were removed due to different dates. In addition, MERIS or MODIS data were not applied in the study as clouded.

To preprocess the data, the hyperspectral reflectance and polarimetric SAR data were analyzed statistically, and outliers with unusual values were identified. Reflectance between the red and NIR range was considered, and if the reflectance at a certain band was particularly high or extremely low compared with that for other stations, this station was treated as an outlier. Then, data from 75 stations were selected for this study. The spectral reflectance of the water samples from the 75 stations is shown in Fig. 2, where we can see a typical pattern of the optical reflectance of turbid productive waters. Fig. 3 shows the HH and VV polarization SAR data of the 75 stations in order of station number. In the study, 38 stations were used to calibrate the chl-a estimation models, and the other 37 stations were used as testing data to validate the effectiveness of the models. The calibration and validation stations were selected randomly, considering their geographic locations.

C. Methods

Two semiempirical models (i.e., two-band and three-band models) and two empirical methods (i.e., MLR and MLP) were employed to estimate chl-a concentration in Lake Taihu. The semiempirical models were proposed based on the optical reflectance of chl-a and thus were used with only optical data in the study. Empirical methods are able to estimate chl-a concentration from a statistical approach and thus can be applied to optical data alone, as well as to combined optical and SAR data.

Considering the two data sources and the nature of the methods, different combinations of optical bands and SAR data were designed as the inputs for different methods (see Table I). This study aimed to comprehensively evaluate the following: 1) the potential of HH and VV polarization SAR data for complementing the chl-a estimate, 2) the effectiveness of the two widely used optical models (i.e., two-band and three-band models), and 3) the performance of different empirical approaches in combining the optical and SAR data for chl-a estimation.

1) Two-Band Model: Chl-a concentration in turbid productive waters can be estimated based on the spectral reflectance in

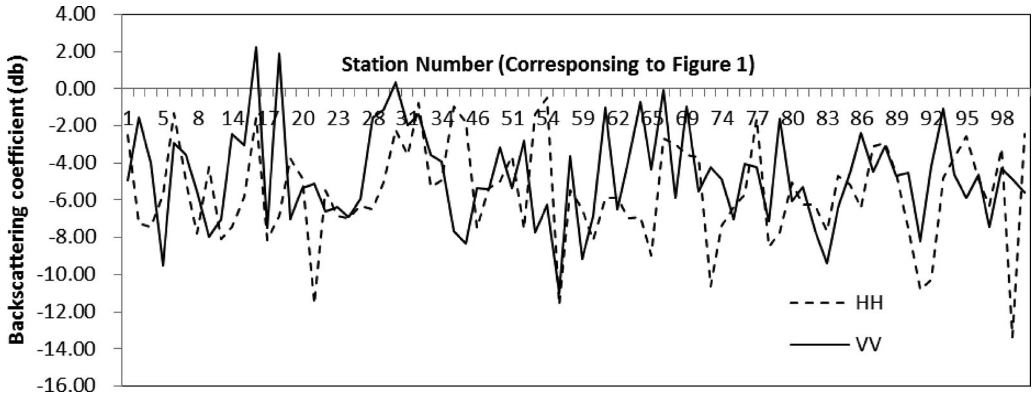

Fig. 3. Backscattering intensity with HH and VV polarizations.

TABLE I
EXPERIMENT DESIGN OF THE CHL-A ESTIMATE ("- - - -" MEANS NO EXPERIMENT)

Methods	2-band	2-band +HH	2-band +VV	2-band +HH+VV	3-band	3-band +HH	3-band +VV	3-band +HH+VV
Two-band model	√	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
Three-band model	- - - -	- - - -	- - - -	- - - -	√	- - - -	- - - -	- - - -
MLR	√	√	√	√	√	√	√	√
MLP	√	√	√	√	√	√	√	√

the red and NIR regions [55]. The widely used two-band model can be formulated as follows [28]:

$$\mathrm{Chl-a} \propto \frac{R(\lambda_{\mathrm{NIR}})}{R(\lambda_{\mathrm{red}})} \tag{1}$$

where Chl-a denotes the chl-a concentration, $R(\lambda_{\mathrm{NIR}})$ is the reflectance in the NIR spectral region, and $R(\lambda_{\mathrm{red}})$ is the reflectance in the red spectral region. The two-band model was employed by various studies, and different spectral bands in the red and NIR regions were used in different areas. The optimal position of the red band was reported to be around 660–670 nm, and that of NIR band was between 700 and 710 nm [26], [56], [57]. In this study, the red band was selected at 665 nm, and the NIR band was set at 705 nm [17].

2) Three-Band Model: The well-known three-band model was developed based on the bio-optical characteristics of turbid productive waters [27]. The three bands were located in the NIR–red spectral regions, taking advantage of the absorption of the pigment of interest. The three-band model can be expressed as

$$\mathrm{Chl-a} \propto [R^{-1}(\lambda_1) - R^{-1}(\lambda_2)] \times R(\lambda_3) \tag{2}$$

where Chl-a is the chl-a concentration, $R(\lambda_1)$, $R(\lambda_2)$ and $R(\lambda_3)$ are the reflectance in the three bands λ_1, λ_2, and λ_3, which are located in the NIR–red spectral regions.

The success of the three-band model depends on the exact positioning of the three bands. A stepwise technique, which is actually an iteration process, has been widely adopted to search the optimal positions of the three bands [25]–[27], [38]. Nevertheless, the optimal positions of the three bands were identified to depend on different study waters [17], [27] and the optimal wavelengths were determined to be 660–670 nm for λ_1, 700–730 nm for λ_2, and 740–750 nm for λ_3 in the waters in Eastern Nebraska, USA. In Chesapeake Bay, the three bands were at 674–676 nm (λ_1), 691–698 nm (λ_2), and 723–739 nm (λ_3) [29]. Referring to this study in Taihu Lake, China, the optimal positions of the three bands were reported at 660–675 nm for λ_1, 700–710 nm for λ_2, and 730–750 for λ_3 [38]. However, according to the original meaning of the third band, λ_3 should be minimally affected by the absorption of chl-a [27]. In this study, the three bands were optimized with a stepwise process to be finally at 665 nm (λ_1), 705 nm (λ_2), and 730 nm (λ_3).

3) Multiple Linear Regression: MLR can be used to empirically estimate chl-a concentration using both optical and SAR data. The simple idea of the MLR is to search for an optimal linear combination of a set of independent variables [58], [59]. In the study of chl-a estimation, this linear combination can be formulated with the following [60], [61]:

$$\mathrm{Chl-a} = \beta_0 + \sum_{i=1}^{N} \beta_i R(\lambda_i) \tag{3}$$

where Chl-a is the chl-a concentration, β_0 and β_i are the constant coefficients that will be determined by the least-squared method [62], and $R(\lambda_i)$ is the reflectance of the band λ_i. Especially, in this study, the two polarization SAR coefficients are treated as two bands when combining the optical and SAR data. Commonly, a stepwise regression process can be run to test the independency of each variable, and to select the most related variables for the optimal regression [58], [59]. However, as the optical bands have been selected with another approach described earlier, there is no further stepwise procedure conducted in this study.

4) Multilayer Perceptron Network: Chl-a estimation using both optical and SAR data can also employ the MLP, which is structured with three types of layers: input, hidden, and output. For the application of chl-a estimation, the input layer corresponds to the bands of both optical and polarimetric data. The MLP treats each of the inputs (i.e., same optical bands plus SAR VV and/or HH data) independently and runs a nonlinear regression to estimate the chl-a concentration. Operations relating to regression are conducted within the hidden layer (s), depending on both the number of hidden layers and the number of nodes in each hidden layer. Basic knowledge about MLP has been described in detail in many studies [63].

The first factor is the number of hidden layers. One or two hidden layers may be involved in the application of MLP in remote-sensing studies [64]. Another key factor is the number of nodes. Various methods have been discussed to determine the number of nodes for each hidden layer, but none is widely accepted [65]. Weng and Hu [64] estimated the number of nodes in the hidden layer as follows:

$$N_h = \mathrm{INT}\sqrt{N_i \times N_o} \tag{4}$$

where N_h denotes the number of nodes in the hidden layer, N_i is the number of nodes in the input layer and N_o is the number of nodes in the output layer and INT stands for rounding operation. In the regression case of chl-a estimation, N_o is treated as 1.

When the basic structure of the MLP is set, its weight can be adjusted through the training process with the labeled training data set. The target of the training process is minimizing the error between the actual label value and the predicted output by the MLP. The back-propagating algorithm is used to effectively search for the weights in the MLP, in which the gradient of the error is propagated from the output to the input layer. Once the MLP is trained, the knowledge about the data samples is contained in the weights of all the nodes on the MLP. To control the searching step of optimized weights and help to accelerate the searching process, parameters of learning rate and momentum should be determined empirically [65].

D. Accuracy Assessment

To evaluate the performance of each experimental approach, the coefficient of determination (R-squared) and mean absolute error (MAE) are employed to calculate the accuracy of the chl-a estimation. R-squared and MAE are calculated with the following:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}(|\mathrm{CHL\_}M_i - \mathrm{CHL\_}P_i|) \tag{5}$$

$$R - \mathrm{Squared} = 1 - \frac{SS_{\mathrm{err}}}{SS_{\mathrm{tot}}} \tag{6}$$

$$SS_{\mathrm{err}} = \sum_{i=1}^{n}\left(\mathrm{CHL\_}M_i - \mathrm{CHL\_}P_i\right)^2 \tag{7}$$

$$SS_{\mathrm{tot}} = \sum_{i=1}^{n}\left(\mathrm{CHL\_}M_i - \overline{\mathrm{CHL\_}M}\right)^2 \tag{8}$$

$$\overline{\mathrm{CHL\_}M} = \frac{1}{n}\sum_{i=1}^{n}\mathrm{CHL\_}M_i. \tag{9}$$

In the above equations, n is the number of testing water sampling points, CHL_M_i denotes the measured chl-a concentration, and CHL_P_i is the predicted chl-a concentration.

III. Results and Discussion

A. Chl-a Estimation With Optical Data

With only optical data, the two-band model, three-band model, MLR, and MLP were employed to estimate chl-a concentration. Fig. 4 depicts the results of the estimation with the accuracy of R-squared and MAE. Some interesting results can be observed in Fig. 4:

1) The 37 predicted points can be divided into two groups for all the estimation models: points with chl-a lower than 26 ug/L and chl-a higher than 26 ug/L. A general underestimation can be observed for chl-a concentration above 26 ug/L, and this underestimation is even greater when using MLR and MLP.
2) The best accuracy comes from the three bands with MLP, obtaining an R-squared value of 0.70 and a MAE of 9.44 ug/L. The three-band model produced low accuracy (R-squared: 0.45; MAE: 8.24 ug/L), which is similar to those in other studies [2], [16], [17]. From Fig. 4(b), we can see that the main cause of this low accuracy lies in an overestimated point, where the predicted chl-a is 54.94 ug/L and the measured chl-a is 31.24 ug/L.
3) Considering the methods, MLR is more suitable for chl-a in lower concentrations, as MLP seems to generally underestimate chl-a with low concentrations (<26 ug/L in this study). In contrast, MLP appears more appropriate for high chl-a concentrations than the corresponding cases in MLR [see Fig. 4(c) and (e) and (d) and (f)].

B. Chl-a Estimate With Optical and Polarimetric SAR Data

MLR and MLP were employed to combine the optical and polarization SAR data. During this process, the two polarization SAR backscattering coefficients were treated as two bands to be combined with the optical data. To evaluate the usefulness of the VV polarization information, the VV SAR data were combined with the two bands and three bands. In this way, the results from MLR and MLP can be compared with those from the two-band and three-band models. Fig. 5 shows the results of the chl-a estimation. It indicates that VV polarization data may add little positive information when using three bands [23], [25]. For instance, the R-squared remained 0.70 ug/L when using the MLP, but MAE decreased to 7.41 ug/L. However, this improvement was still unstable because the accuracy decreased slightly with the MLR.

To further investigate the potential of the VV polarization data, the total 75 points were sorted in increasing order of chl-a concentration. Then, combined with the corresponding VV backscattering coefficient, the data were plotted in Fig. 6. It is a slight decrease in the VV backscattering coefficient with the

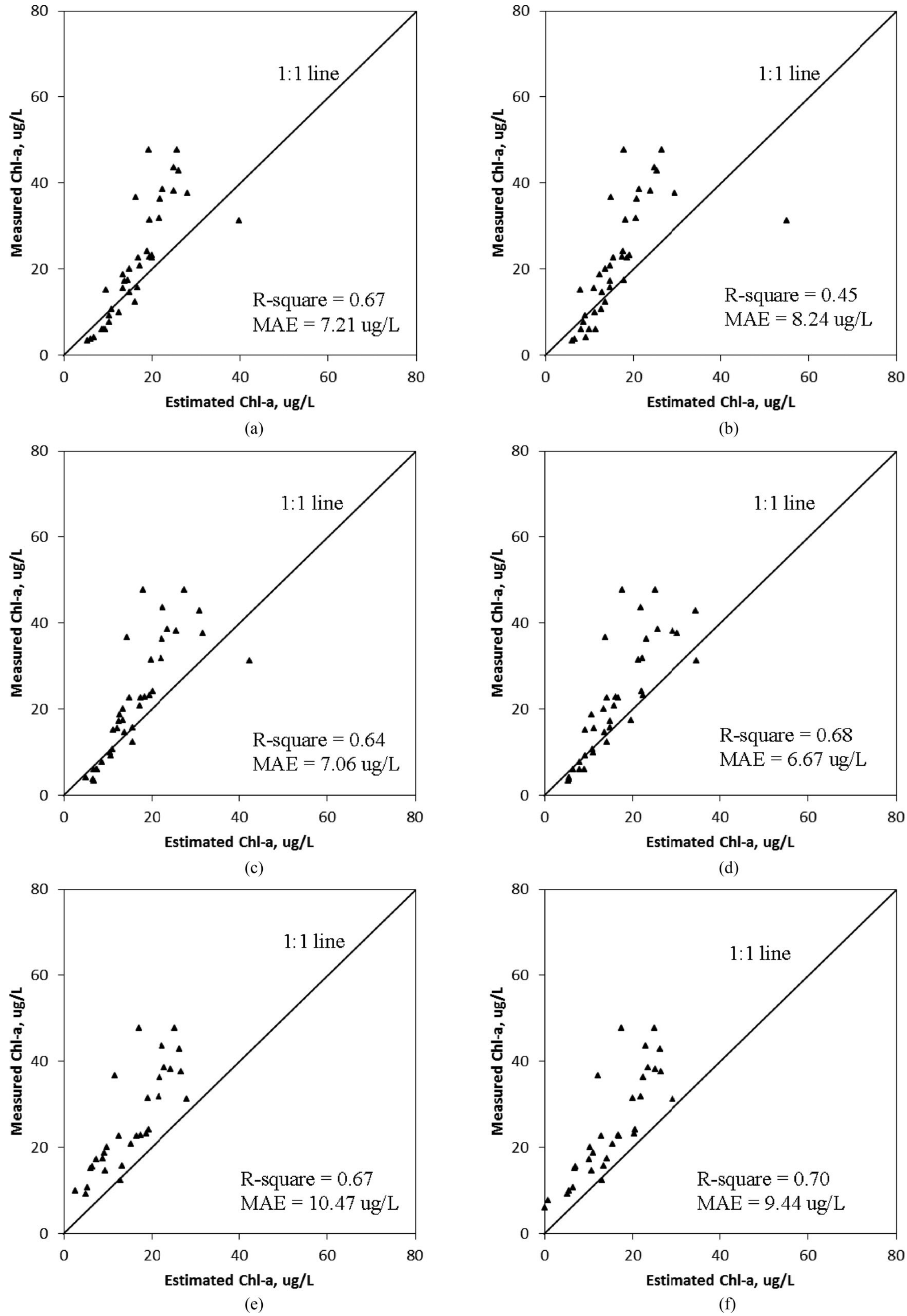

Fig. 4. Chl-a estimation with optical data. (a) Two-band model. (b) Three-band model. (c) Two bands MLR. (d) Three bands MLR. (e) Two bands MLP. (f) Three bands MLP.

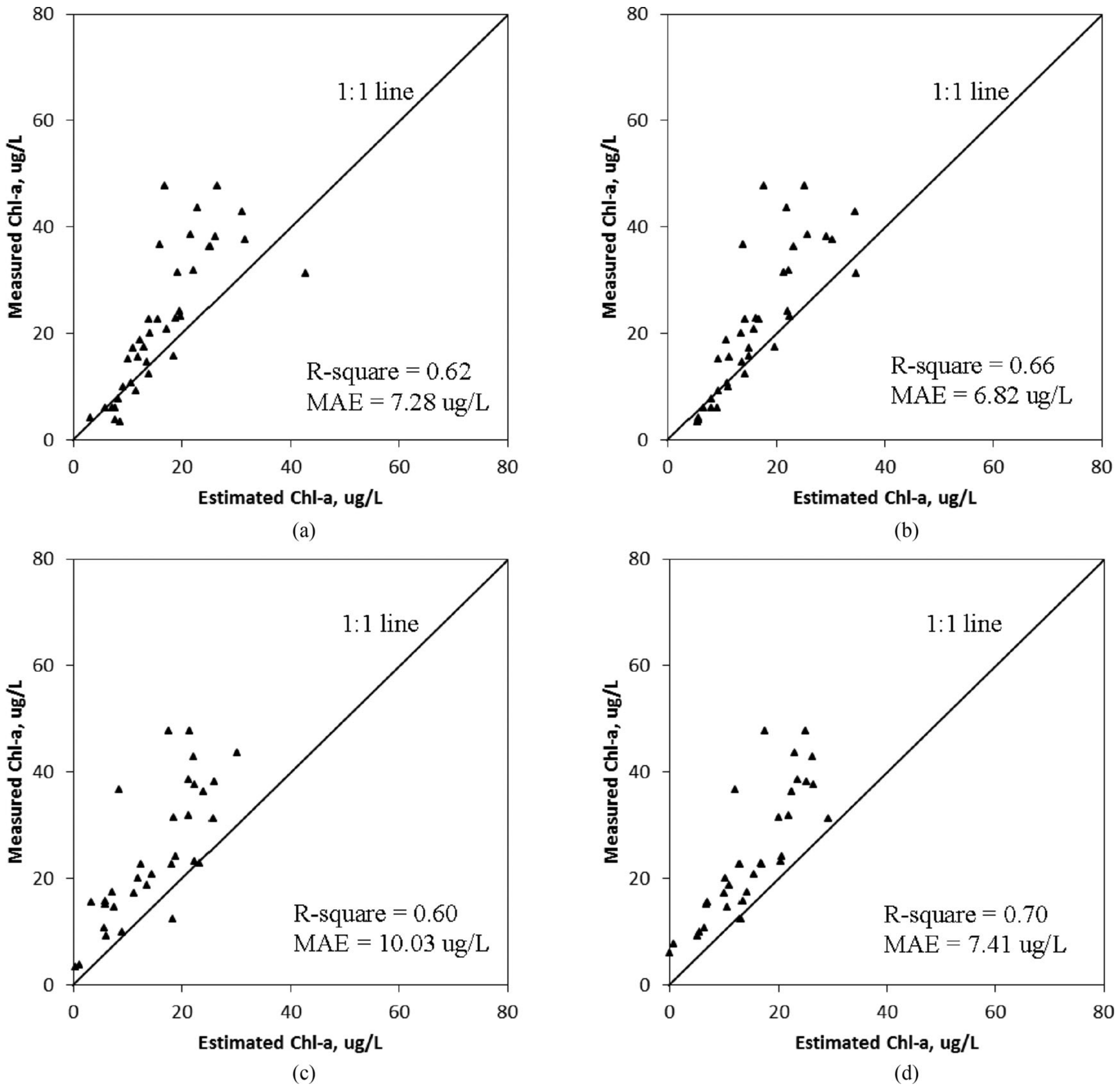

Fig. 5. Chl-a estimation with optical data and SAR VV data. (a) Two bands + VV (MLP). (b) Three bands + VV (MLP). (c) Two bands + VV (MLP). (d) Three bands + VV (MLP).

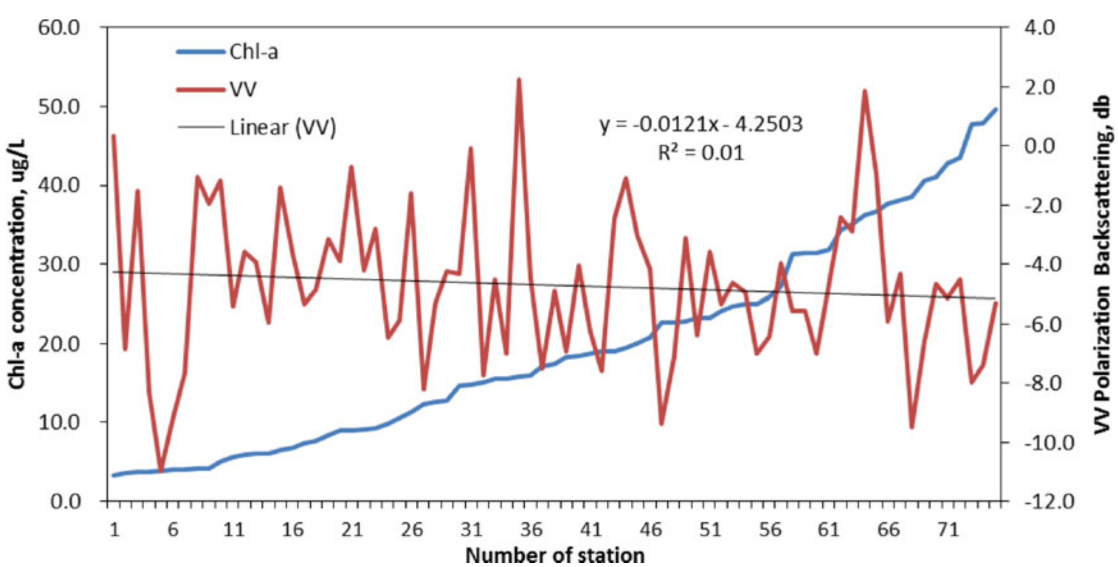

Fig. 6. Trend of VV polarization SAR data versus measured chl-a concentration.

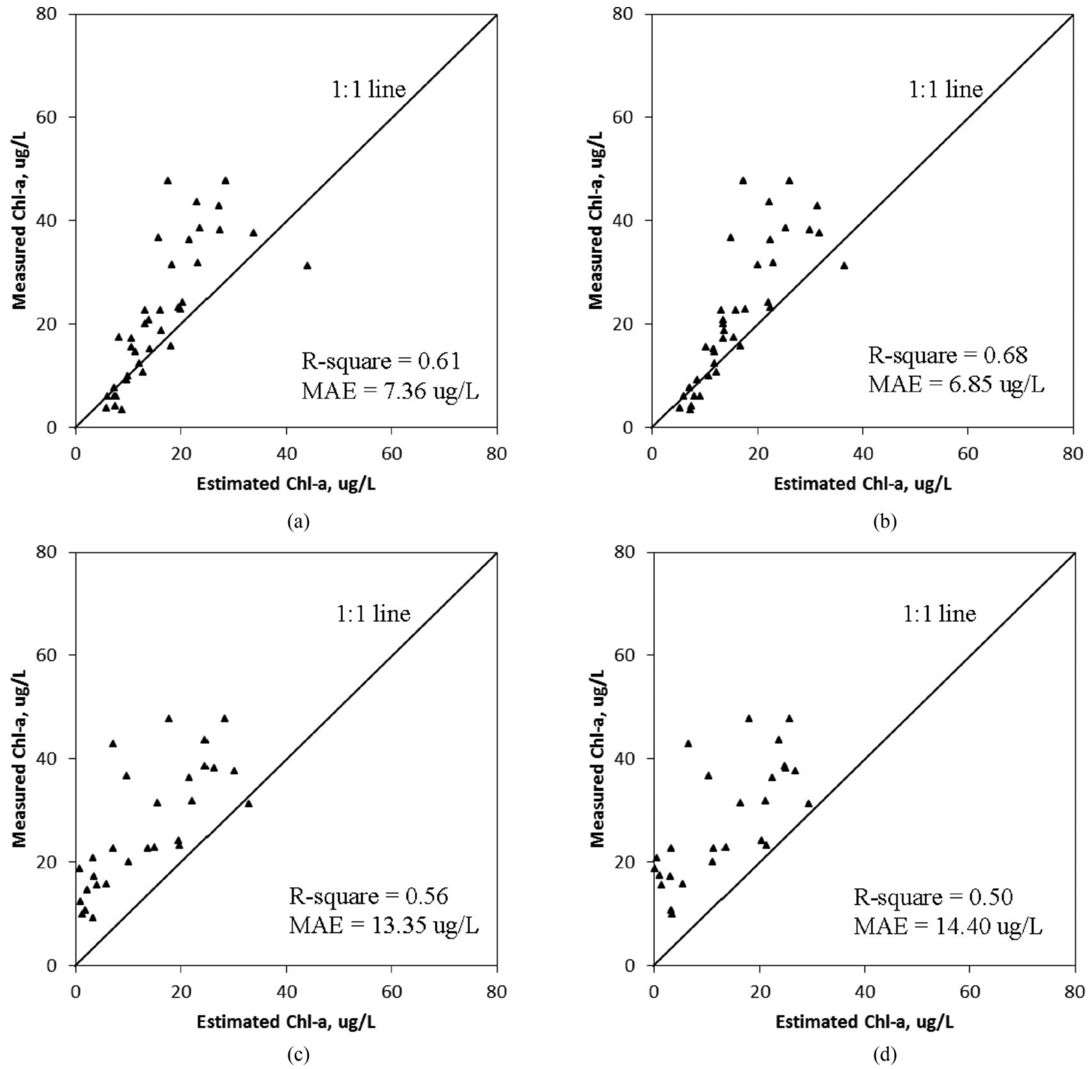

Fig. 7. Chl-a estimation with optical and SAR HH data. (a) Two bands + HH (MLR). (b) Three bands + HH (MLR). (c) Two bands + HH (MLP). (d) Three bands + HH (MLP).

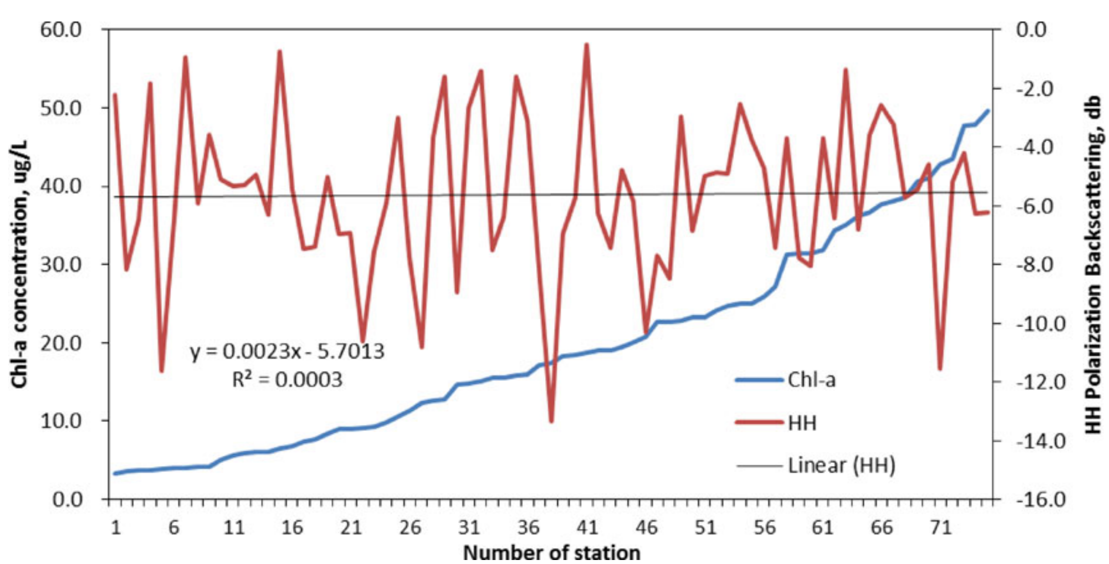

Fig. 8. Trend of HH polarization SAR data versus measured chl-a concentration.

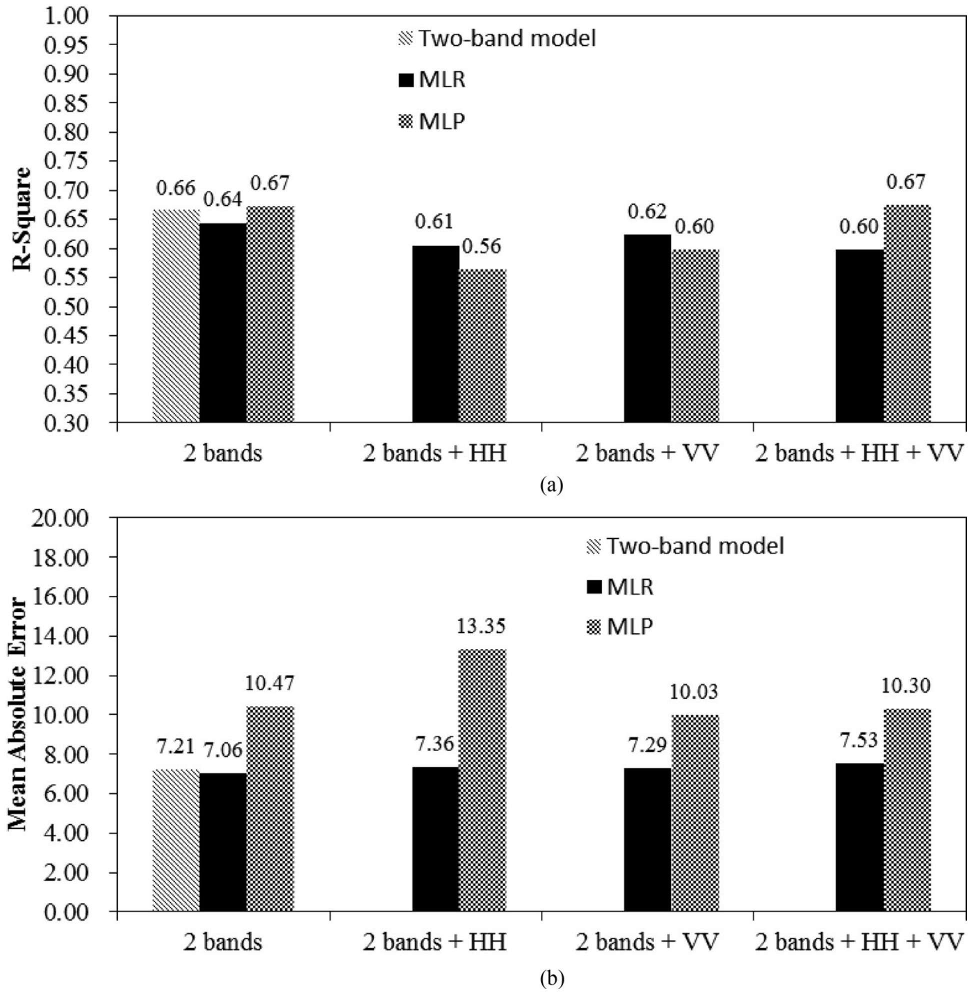

Fig. 9. Retrieval accuracy using combinations of two bands and SAR data. (a) R-square. (b) MAE.

increase in the chl-a concentration [33], [36], so the decreasing trend is not statistically significant, which is why the accuracy from the combined use of three optical bands and VV is unstable.

On the other hand, the evaluation of the HH polarization data showed a generally negative effect, with a decline in the accuracy for both combinations with two and three optical bands (see Fig. 7). The underestimation became more serious when using the MLP approach, reflected in the low MAE accuracy, so 27 stations data of lower than 26 ug/L were used to validate as shown in Fig. 7(c) and (d). From the trend line of the HH backscattering coefficients with increasing chl-a concentration (see Fig. 8), almost no correlation between HH data and chl-a can be observed. Moreover, the additional use of HH data brought some noise to the estimation models, resulting in a decline in the estimation accuracy. The different effects brought by the VV and HH polarization SAR data may be produced by the water surface roughness and the dielectric constant of the water [36], [39]. However, to investigate this phenomenon, more information should be collected, such as the wind speed and directions, backscattering coefficients of water under different wind situations and chl-a concentration [40]. Unfortunately, due to the limitation of our field experiment, we did not collect such information for analysis. Therefore, here we only mention these possibly different behaviors of polarization SAR data responding to the water surface in Lake Taihu.

C. Comparison of the Chl-a Estimate Using Different Datasets

To compare the results of chl-a estimate with different datasets and approaches and to evaluate the potential of SAR data in a more comprehensive way, we also used dual polarization data. The following comparison experiment was designed in two groups: 1) a combination of two optical bands and SAR data and 2) a combination of three optical bands and SAR data. Four sets of data combination were designed: 1) optical bands only, 2) combined optical bands and HH data, 3) combined optical bands and VV data, and 4) combined optical bands and HH and VV data. R-squared and MAE were calculated as the chl-a estimation accuracy.

Fig. 9 shows the accuracy assessment of combining the two optical bands and SAR data. First, the highest R-squared (0.67) comes from the single use of two bands and the combination of two bands, HH and VV polarization data. Therefore, the additional use of SAR data only reduces the MAE. Second, MLR appears to be the most stable approach, with both a stably medium R-squared (~0.62) and a stably low MAE (~7.31). Third, the

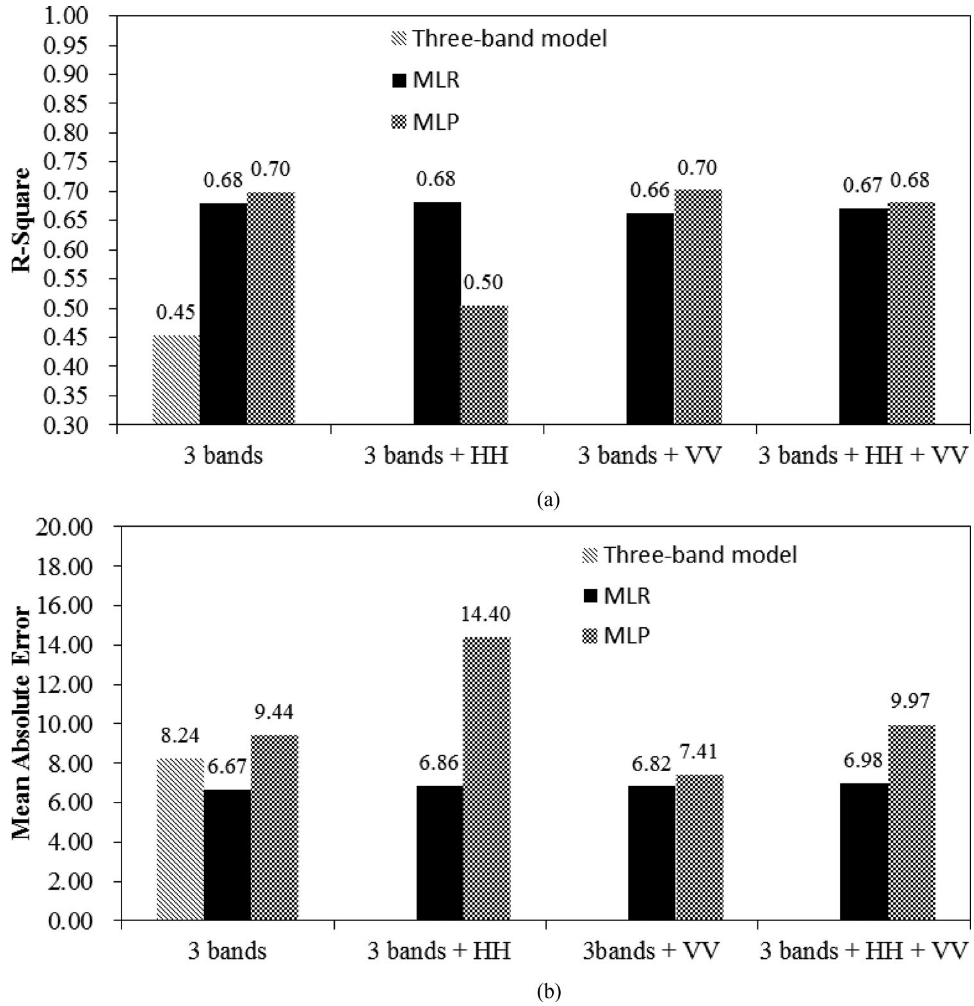

Fig. 10. Retrieval accuracy using combinations of three bands and SAR data. (a) R-square. (b) MAE.

two-band model obtained rather good results compared with other methods and datasets, indicating the effectiveness of the two-band model [25], [27]. The overall low R-squared (<0.67) probably resulted from the large variation range of the chl-a concentration, for which different models may be used for low concentrations of chl-a and high concentrations of chl-a [26], [28].

Accuracy from combining three optical bands and SAR data is illustrated in the bar chart of Fig. 10. First, the best estimation results occurred from both the single use of three bands and the combination of three bands and VV polarization data. The additional use of VV data did not increase R-squared, either, but reduced MAE from 9.44 to 7.41 ug/l. Second, MLR also appeared to be stable for these datasets, with an average R-squared of 0.67 and an average MAE of 6.83 ug/L. Third, what is interesting is that lowest accuracy came from the single use of three optical bands with the three-band model, while the highest accuracy was from the same dataset but with an MLP approach [34], [36]. This result may indicate the unstable nature of the three-band model, for it may produce big errors for some single points Fig. 4(b). However, the use of three optical bands was generally better than using only two, as shown and compared in Figs. 9 and 10.

IV. Conclusion

This study aims to evaluate the potential of polarimetric SAR data as a complementary data source to the hyperspectral data in estimating chl-a. Chl-a estimation has played an increasingly important role in monitoring environmental issues (e.g., algal blooms) in aquatic ecosystems; however, it still represents a great challenge in turbid productive waters where environmental problems are often serious. *In situ* hyperspectral data and one scenic ASAR data for Lake Taihu were collected, and both semiempirical and empirical methods were employed to estimate chl-a concentration. The results show that there is an overall underestimation of chl-a in areas with a concentration higher than 26 ug/L, which is probably caused by a large spatial variation of chl-a in the study area. The results also demonstrate that there is a more stable performance of the MLR model than the MLP, whereas the MLP performs a general underestimation for both low and high chl-a concentrations.

On the other hand, SAR data improved the chl-a estimation slightly in this case study, mostly depending on the concentrations of chl-a, data sets, and estimation models. In the study, the improvement of VV polarization SAR data is better than that of HH polarization data in chl-a estimation with a reduction in

MAE. We still need to apply more SAR and MERIS data in varied turbid waters in the near future to determine how SAR data are useful in improving chl-a estimation.

Acknowledgment

ASAR data from the satellite ground receiving station of the Chinese University of Hong Kong is highly appreciated. Y. Zhang and M. Hallikainen conceived and designed the experiments; H. Zhang performed the experiments; H. Duan and Y. Li analyzed the data; X. S. Liang contributed reagents/materials/analysis tools; Y. Zhang and H. Zhang wrote the paper. The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

[1] J. M. O'Neil, T. W. Davis, M. A. Burford, and C. J. Gobler, "The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change," *Harmful Algae*, vol. 14, pp. 313–334, 2012. doi: 10.1016/j.hal.2011.10.027.

[2] H. T. Duan *et al.*, "Two-Decade reconstruction of Algal Blooms in China's lake Taihu," *Environ. Sci. Technol.*, vol. 43, no. 10, pp. 3522–3528, 2009.

[3] C. M. Hu, Z. P. Lee, R. H. Ma, K. Yu, D. Q. Li, and S. L. Shang, "Moderate resolution imaging spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China," *J. Geophysical Res.-Oceans*, vol. 115, Paper no. 40, 2010.

[4] P. B. O. Ochumba and D. I. Kibaara, "Observations on blue-green-Algal blooms in the open waters of lake victoria, Kenya," *African J. Ecology*, vol. 27, no. 1, pp. 23–34, 1989. doi: 10.1111/j.1365-2028.1989.tb00925.x.

[5] L. Sitoki, J. Gichuki, C. Ezekiel, F. Wanda, O. C. Mkumbo, and B. E. Marshall, "The environment of lake victoria (East Africa): Current status and historical changes," *Int. Rev. Hydrobiol.*, vol. 95, no. 3, pp. 209–223, 2010.

[6] M. Tijdens *et al.*, "Population dynamics and diversity of viruses, bacteria and phytoplankton in a shallow eutrophic lake," *Microbial Ecology*, vol. 56, no. 1, pp. 29–42, 2008.

[7] R. K. Vincent *et al.*, "Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie," *Remote Sensing Environ.*, vol. 89, no. 3, pp. 381–392, 2004.

[8] J. Luo *et al.*, "Mapping species of submerged aquatic vegetation with multi-seasonal satellite images and considering life history information," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 57, pp. 154–165, 2017.

[9] I. Chorus and J. Bartram, *Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring, and Management*. New York, NY, USA: Taylor & Francis, 1999.

[10] H. W. Paerl *et al.*, "Controlling harmful cyanobacterial blooms in a hypereutrophic lake (Lake Taihu, China): The need for a dual nutrient (N & P) management strategy," *Water Res.*, vol. 45, no. 5, pp. 1973–1983, 2011.

[11] L. C. Backer, "Cyanobacterial harmful algal blooms (CyanoHABs): Developing a public health response," *Lake Reservoir Manage.*, vol. 18, no. 1, pp. 20–31, 2002.

[12] H. T. Duan, R. H. Ma, J. P. Xu, Y. Z. Zhang, and B. Zhang, "Comparison of different semi-empirical algorithms to estimate chlorophyll-a concentration in inland lake water," *Environ. Monitoring Assessment*, vol. 170, nos. 1–4, pp. 231–244, 2010.

[13] Li, L., R. E. Sengpiel, D. L. Pascual, L. P. Tedesco, J. S. Wilson, and E. Soyeux, "Using hyperspectral remote sensing to estimate chlorophyll-a and phycocyanin in a mesotrophic reservoir," *Int. J. Remote Sensing*, vol. 31, no. 15, pp. 4147–4162, 2010.

[14] S. G. H. Simis, S. W. M. Peters, and H. J. Gons, "Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water," *Limnology Oceanography*, vol. 50, no. 1, pp. 237–245, 2005.

[15] Y. P. Wang, H. Xia, J. M. Fu, and G. Y. Sheng, "Water quality change in reservoirs of Shenzhen, China: detection using LANDSAT/TM data," *Sci. Total Environ.*, vol. 328, nos. 1–3, pp. 195–206, 2004.

[16] H. T. Duan, S. A. Loiselle, L. Zhu, L. Feng, Y. Zhang, and R. Ma, "Distribution and incidence of algal blooms in Lake Taihu," *Aquatic Sci.*, vol. 77, pp. 9–16, 2015.

[17] H. T. Duan *et al.*, "A new three-band algorithm for estimating chlorophyll concentrations in turbid inland lakes," *Environ. Res. Lett.*, vol. 5, no. 4, 2010, Art. no. 044009.

[18] H. T. Duan, R. H. Ma, and C. M. Hu, "Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of East China," *Remote Sensing Environ.*, vol. 126, pp. 126–135, 2012.

[19] H. J. Gons, M. Rijkeboer, and K. G. Ruddick, "A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters," *J. Plankton Res.*, vol. 24, no. 9, pp. 947–951, 2002.

[20] T. Kutser, L. Metsamaa, N. Strombeck, and E. Vahtmae, "Monitoring cyanobacterial blooms by satellite remote sensing," *Estuarine Coastal Shelf Sci.*, vol. 67, nos. 1/2, pp. 303–312, 2006.

[21] A. Y. Morel and H. R. Gordon, "Report of the working group on water color," *Boundary-Layer Meteorology*, vol. 18, no. 3, pp. 343–355, 1980.

[22] V. E. Brando and A. G. Dekker, "Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1378–1387, Jun. 2003.

[23] A. Gitelson, "The Peak near 700 Nm on radiance spectra of Algae and water - relationships of its magnitude and position with chlorophyll concentration," *Int. J. Remote Sens.*, vol. 13, no. 17, pp. 3367–3373, 1992.

[24] D. Gurlin, A. A. Gitelson, and W. J. Moses, "Remote estimation of chl-a concentration in turbid productive waters - Return to a simple two-band NIR-red model?," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3479–3490, 2011.

[25] A. A. Gitelson *et al.*, "A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation," *Remote Sens. Environ.*, vol. 112, no. 9, pp. 3582–3593, 2008.

[26] G. Dall'Olmo and A. A. Gitelson, "Effect of bio-optical parameter variability on the remote estimation of chlorophyll-a concentration in turbid productive waters: Experimental results," *Appl. Opt.*, vol. 44, no. 3, pp. 412–422, 2005.

[27] G. Dall'Olmo, A. A. Gitelson, and D. C. Rundquist, "Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters," *Geophysical Res. Letters*, vol. 30, no. 18, pp. 159–171, 2003.

[28] A. A. Gitelson and K. Y. Kondratyev, "Optical-models of mesotrophic and Eutrophic water bodies," *Int. J. Remote Sensing*, vol. 12, no. 3, pp. 373–385, 1991.

[29] A. A. Gitelson, J. F. Schalles, and C. M. Hladik, "Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study," *Remote Sens. Environ.*, vol. 109, no. 4, pp. 464–472, 2007.

[30] H. J. Gons, M. T. Auer, and S. W. Effler, "MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes," *Remote Sens. Environ.*, vol. 112, no. 11, pp. 4098–4106, 2008.

[31] L. H. Han and D. C. Rundquist, "Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: A case study in a turbid reservoir," *Remote Sens. Environ.*, vol. 62, no. 3, pp. 253–261, 1997.

[32] D. Odermatt, A. Gitelson, V. E. Brando, and M. Schaepman, "Review of constituent retrieval in optically deep and complex waters from satellite imagery," *Remote Sens. Environ.*, vol. 118, pp. 116–126, 2012.

[33] L. Gross, S. Thiria, R. Frouin, and B. G. Mitchell, "Artificial neural networks for modeling the transfer function between marine reflectance and phytoplankton pigment concentration," *J. Geophysical Res.-Oceans*, vol. 105, no. C2, pp. 3483–3495, 2000.

[34] L. E. Keiner and X. H. Yan, "A neural network model for estimating sea surface chlorophyll and sediments from thematic mapper imagery," *Remote Sens. Environ.*, vol. 66, no. 2, pp. 153–165, 1998.

[35] D. Y. Sun, Y. M. Li, and Q. Wang, "A unified model for remotely estimating chlorophyll a in lake Taihu, China, based on SVM and in situ hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2957–2965, Aug. 2009.

[36] Y. Z. Zhang, J. Pulliainen, S. Koponen, and M. Hallikainen, "Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data," *Remote Sens. Environ.*, vol. 81, nos. 2/3, pp. 327–336, Aug. 2002.

[37] Y. Z. Zhang, H. Lin, C. Q. Chen, L. D. Chen, B. Zhang, and A. A. Gitelson, "Estimation of chlorophyll-a concentration in estuarine waters: case study of the Pearl River estuary, South China Sea," *Environ. Res. Lett.*, vol. 6, no. 2, pp. 24016-1–24016-9, 2011.

[38] Y. L. Zhang, M. L. Liu, B. Q. Qin, H. J. van der Woerd, J. S. Li, and Y. L. Li, "Modeling remote-sensing reflectance and retrieving chlorphyll-a concentration in extremely turbid case-2 waters (lake Taihu, China)," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 1937–1948, Jul. 2009.

[39] H. S. Zhang, Y. Z. Zhang, and H. Lin, "A comparison study of impervious surfaces estimation using optical and SAR remote sensing images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 18, pp. 148–156, 2012.

[40] F. M. Henderson and Z. G. Xia, "SAR applications in human settlement detection, population estimation and urban land use pattern analysis: A status report," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 79–85, Jan. 1997.

[41] C. Tison, J. M. Nicolas, F. Tupin, and H. Maitre, "A new statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, Oct. 2004.

[42] Y. Z. Zhang, J. Pulliainen, S. Koponen, and M. Hallikainen, "Empirical algorithms for Secchi disk depth using optical and microwave remote sensing data from the Gulf of Finland and the Archipelago Sea," *Boreal Environ. Res.*, vol. 8, no. 3, pp. 251–261, 2003.

[43] Y. Z. Zhang, J. T. Pulliainen, S. S. Koponen, and M. T. Hallikainen, "Water quality retrievals from combined Landsat TM data and ERS-2 SAR data, in the Gulf of Finland," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 622–629, Mar. 2003.

[44] S. Bhattacharya Manickam, G. Singh, and Y. Yamaguchi, "Estimation of snow surface dielectric constant from polarimetric SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 1, pp. 211–218, Jan. 2017.

[45] Y. Alvarez, B. Gonzalez-Valdes, J. A. Martinez-Lorenzo, C. M. Rappaport, and F. Las-Herras, "SAR imaging-based techniques for low permittivity lossless dielectric bodies characterization," *IEEE Antennas Propag. Mag.*, vol. 57, no. 2, pp. 267–276, Apr. 2015.

[46] G. Wang, J. Li, B. Zhang, Q. Shen, and F. Zhang, "Monitoring cyanobacteria-dominant algal blooms in eutrophicated Taihu Lake in China with synthetic aperture radar images," *Chin. J. Oceanology Limnology*, vol. 33, no. 1, pp. 139–148, 2015.

[47] G. Wang, J. Li, B. Zhang, Z. Cai, F. Zhang, and Q. Shen, "Synthetic aperture radar detection and characteristic analysis of cyanobacterial scum in Lake Taihu," *J. Appl. Remote Sensing*, vol. 11, no. 1, 2017, Art. no. 012006.

[48] B. Qin, *Lake Taihu, China: Dynamics and Environmental Change (Monographiae Biologicae)*: Amsterdam, The Netherlands: Springer, 2008.

[49] L. Guo, "Ecology - Doing battle with the green monster of Taihu Lake," *Science*, vol. 317, no. 5842, pp. 1166, 2007.

[50] H. Xu, H. W. Paerl, B. Q. Qin, G. W. Zhu, and G. Gao, "Nitrogen and phosphorus inputs control phytoplankton growth in eutrophic Lake Taihu, China," *Limnology Oceanography*, vol. 55, no. 1, pp. 420–432, 2010.

[51] C. F. Le, Y. M. Li, Y. Zha, D. Y. Sun, C. C. Huang, and H. Lu, "A four-band semi-analytical model for estimating chlorophyll a in highly turbid lakes: The case of Taihu Lake, China," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1175–1182, 2009.

[52] Y. M. Li *et al.*, "Estimation of chlorophyll a concentration using NIR/Red Bands of MERIS and classification procedure in inland Turbid water," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 988–997, Mar. 2012.

[53] J. L. Mueller and G. S. Fargion, *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*, Revision 6. Greenbelt, MD, USA: Goddard Space Flight Center, 2003.

[54] [Online]. Available: http://www.gps.gov/systems/gps/performance/accuracy/

[55] R. P. Stumpf and M. A. Tyler, "Satellite detection of bloom and pigment distributions in estuaries," *Remote Sens. Environ.*, vol. 24, no. 3, pp. 385–404, 1988.

[56] F. E. Hoge, C. W. Wright, and R. N. Swift, "Radiance-ratio algorithm wavelengths for remote oceanic chlorophyll determination," *Appl. Opt.*, vol. 26, no. 11, pp. 2082–2094, 1987.

[57] Y. Z. Yacobi, A. Gitelson, and M. Mayo, "Remote-sensing of chlorophyll in lake kinneret using high-spectral-resolution radiometer and landsat-tm - spectral features of reflectance and algorithm development," *J. Plankton Res.*, vol. 17, no. 11, pp. 2155–2173, 1995.

[58] I. H. Bernstein, C. P. Garbin, and G. K. Teng, *Applied Multivariate Analysis*. New York, NY, USA: Springer-Verlag, 1988.

[59] R. F. Kokaly and R. N. Clark, "Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression," *Remote Sens. Environ.*, vol. 67, no. 3, pp. 267–287, 1999.

[60] H. Camdevyren, N. Demyr, A. Kanik, and S. Keskyn, "Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs," *Ecological Modelling*, vol. 181, no. 4, pp. 581–589, 2005.

[61] N. R. Draper and H. Smith, *Applied Regression Analysis*, Wiley Series in Probability and Statistics Texts and References Section, 3rd ed. New York, NY, USA: Wiley, 1998.

[62] J. D. Carroll, P. E. Green, and A. Chaturvedi, *Mathematical Tools for Applied Multivariate Analysis*, Rev. ed. San Diego, CA, USA: Academic, 1997.

[63] G. D. Garson, "Neural networks: An introductory guide for social scientists," *J. Artificial Societies Social Simulation*, vol. 4, no. 3, pp. 321–322, 2001.

[64] Q. H. Weng and X. F. Hu, "Medium spatial resolution satellite imagery for estimating and mapping urban impervious surfaces using LSMA and ANN," *IEEE Trans. Geoscience Remote Sensing*, vol. 46, no. 8, pp. 2397–2406, Aug. 2008.

[65] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed. Berlin, Germany: Springer, 2006.

Authors' photographs and biographies not available at the time of publication.