Deep-Learning-Based Retrieval of an Orange Band Sensitive to Cyanobacteria for Landsat-8/9 and Sentinel-2

Milad Niroumand-Jadidi, *Member, IEEE*, and Francesca Bovolo, *Senior Member, IEEE*

***Abstract*—The lack of an orange band (∼620 nm) in the imagery captured by Landsat-8/9 and Sentinel-2 restricts the detection and quantification of harmful cyanobacterial blooms in inland waters. A recent study suggested the retrieval of orange remote sensing reflectance, R_{rs} (620), by assuming green, red, and panchromatic (Pan) bands of Landsat-8 as predictors through a linear model. However, this method is not applicable to Sentinel-2 imagery lacking a Pan band. Moreover, the Pan-based method does not account for the nonlinear relationships among the R_{rs} data at different wavelengths. We propose a deep-learning model called Deep OrAnge Band LEarning Network (DOABLE-Net) that leverages a large training set of R_{rs} data from radiative transfer simulations and in situ measurements. The proposed DOABLE-Net is structured as five fully connected layers and implemented either with or without the Pan band as an input feature, which only the latter applies to Sentinel-2. DOABLE-Net provided more accurate and robust retrievals than the Pan-based method on a wide range of independent validation datasets. The performance of DOABLE-Net on Landsat-8/9 data was minimally impacted by including the Pan band. The results from Sentinel-2 data analysis also confirmed that the DOABLE-Net provides promising results without using a Pan band.**

***Index Terms*—Cyanobacteria, deep learning, inland waters, Landsat-8/9, orange band, phycocyanin, Sentinel-2.**

I. Introduction

REMOTE sensing of water quality in inland and nearshore coastal waters has undergone significant progress over the past decade with the launch of Landsat-8/9 and Sentinel-2 satellites [1], [2]. Operational land imager (OLI) and multispectral imager (MSI) sensors aboard Landsat-8/9 and Sentinel-2, respectively, are not originally designed for aquatic applications. However, their spatial resolution (OLI 30 m, MSI 10−60 m) and dynamic range (12−14 bit) have been favorable in various studies conducted in optically complex inland and coastal waters [3], [4], [5]. OLI and MSI capture five and seven bands, respectively, within the visible and near-infrared portion of the spectrum (< 800 nm), which are useful for water quality studies. Although OLI and MSI opened up new opportunities, their spectral band designations are not optimal for retrieving some water constituents. In particular, mapping and quantifying cyanobacteria have appeared challenging due to the lack of an orange band in these sensors [6].

Cyanobacterial harmful algal blooms (cyano-HABs) can develop dense biomasses producing toxins that severely endanger the aquatic habitat and associated ecosystem services [7], [8]. Given the substantial adverse effects of cyano-HABs, the accurate and timely monitoring of cyanobacteria biomass by means of satellite remote sensing is a high priority. Most ocean color sensors, such as ocean and land color instrument onboard Sentinel-3, capture an orange band centered at ∼620 nm that facilitates the detection of phycocyanin, which is the characteristic pigment of cyanobacteria [9], [10]. However, the spatial resolution of these sensors (hundreds of meters) is too coarse for studying inland waters. On the other hand, the acquisition of images from spaceborne hyperspectral missions, such as PRISMA [11], [12], is not regular despite the sufficient spatial resolution (30 m). Therefore, Landsat-8/9 and Sentinel-2 remain the key satellites for studying inland and coastal water quality, offering suitable spatial and temporal resolutions (5–8 days).

Phycocyanin is a photosynthetic pigment with a unique spectral response of an absorption peak at ∼ 620 nm and a reflectance peak at ∼650 nm that distinguishes cyanobacteria from other phytoplankton communities [13], [14]. However, OLI and MSI lack such an orange band that undermines their utility for mapping cyanobacteria. A recent study proposed an approach to retrieve the remote sensing reflectance at 620 nm, $R_{rs}(620)$, for Landsat-8 imagery based on a linear relation incorporating the panchromatic (Pan), green, and red bands [9]. Note that R_{rs} is the ratio of water-leaving radiance to downwelling irradiance just above the surface and has units of inverse steradians (sr^{-1}) [15]. The Pan-based orange band retrieval technique provided promising results in mapping phycocyanin concentration in Lake Erie [16] and Eagle Creek Reservoir [17] and has been implemented in the ACOLITE processor [18]. The main hypothesis of this method is to exploit the spectral information available in the panchromatic band, which is a wide band overlapping with multispectral channels. However, this method, called the Pan-based method hereafter, is built upon a simple linear model that employs Pan, green, and red bands as predictors for the orange band retrieval. Thus, it does not account for the complex nonlinear relations among R_{rs} data at different wavelengths, which are typical in optically complex waters [19]. Moreover,

Manuscript received 11 January 2023; revised 24 March 2023; accepted 11 April 2023. Date of publication 13 April 2023; date of current version 28 April 2023. *(Corresponding author: Milad Niroumand-Jadidi.)*

The authors are with the Digital Society Center, Fondazione Bruno Kessler, 38123 Trento, Italy (e-mail: mniroumand@fbk.eu; bovolo@fbk.eu).

Digital Object Identifier 10.1109/JSTARS.2023.3266929

the Pan-based method is not applicable to the imagery from sensors, such as MSI, which does not capture a panchromatic band.

This study aims to develop an advanced deep-learning-based model called Deep OrAnge Band Learning Network (DOABLE-Net) for retrieving a virtual orange band from multispectral bands of Landsat-8/9 or Sentinel-2. Our proposed method is grounded in the high level of spectral interdependence of R_{rs} data [20], [21], [22] and the ability of deep models to learn complex nonlinear features from a large set of training data. In a recent study, a deep-learning model is developed for retrieving near-blue ultraviolet R_{rs} bands from the visible ones obtained by ocean color sensors and provided encouraging results [19]. We build upon this new line of research to advance the state-of-the-art in orange band retrieval by pursuing the following objectives.

1) Develop DOABLE-Net for Landsat-8/9 and Sentinel-2 with high generalization capability by leveraging a large training dataset of radiative transfer simulations and in situ spectral measurements.
2) Examine the impact of involving the Landsat-8/9 Pan band on the performance of the proposed DOABLE-Net.
3) Evaluate the accuracy of orange band retrieval in a wide range of bio-optical conditions and perform comparisons with the Pan-based method [9].

The rest of this article is organized as follows. The training and validation datasets are presented in Section II. Section III introduces the proposed DOABLE-Net and briefly describes the existing Pan-based method and its calibration for retrieving the orange band from Landsat-8/9 data. Section IV illustrates the validation results of DOABLE-Net compared with the Pan-based method. Finally, Section V concludes this article.

II. Datasets

One of the key aspects of developing supervised deep-learning models is creating a large and representative training dataset. In aquatic applications, the in situ data are quite limited in space and time, which hinders building a comprehensive training set. Complementary to in situ data, deep-learning models can largely benefit from radiative transfer simulations [23], [24] providing a massive dataset spanning diverse bio-optical conditions. The radiative transfer models account for the absorption and backscattering properties of pure water and its constituents. The simulation allows the creation of a large set of R_{rs} spectra spanning a wide range of inherent optical properties and concentrations of constituents [23].

We use a large set of hyperspectral R_{rs} for training DOABLE-Net consisting of radiative transfer simulations (100 000 samples) up to 800 nm representing a wide range of bioptical conditions, which are taken from the article presented in [19], and shipborne reflectance measurements (5805 samples) up to 950 nm in the optically complex Baltic Sea conducted by Tilstone et al. [25].

We employ four independent datasets (4311 samples in total) representing very diverse bio-optical conditions to examine the prediction performance of DOABLE-Net. The validation datasets and the sources are summarized in Table I. The data originally used for training and validation of the Pan-based method, acquired at Belgian and Dutch lakes, are also among our validation datasets. The abbreviations of the datasets (see Table I) are taken from the data sources.

TABLE I
IN SITU HYPERSPECTRAL DATASETS USED FOR THE VALIDATION OF DOABLE-NET AND PAN-BASED METHODS

Dataset	#Samples	Description	Source
SpecWA	3685	R_{rs} up to 910 nm at 11 different inland water bodies in Germany spanning very diverse bio-optical conditions	[26]
SeaSWIR	137	R_{rs} up to 1300 nm at turbid estuarine sites (Gironde, La Plata, Scheldt)	[27]
NORCOHAB	44	R_{rs} up 950 nm during North Sea Coast Harmful Algal Bloom	[28]
Belgian-Dutch lakes	445	R_{rs} in diverse bio-optical conditions over several lakes in Belgium (up to 896 nm) and the Netherlands (up to 780 mn)	[29] [9]

Although the four *in situ* R_{rs} datasets provide an independent and comprehensive means of evaluating the performance of orange band retrieval methods, we also apply DOABLE-Net to Landsat-9 (OLI-2) and Sentinel-2 (MSI) imagery from San Francisco Bay (see Fig. 1). The images from the two sensors are acquired on the same day (December 10, 2021). The ACOLITE processor [18], [30] is used to apply dark spectrum fitting atmospheric correction. As shown in Fig. 1, the bay and surrounding coastal area show a wide range of color gradients representing various water types from clear blue to turbid waters with sediment plumes. We perform a consistency analysis between the results of the DOABLE-Net and Pan-based methods, retrieving the $R_{\mathrm{rs}}(620)$ for the Landsat-9 image. Moreover, the consistency of $R_{\mathrm{rs}}(620)$ retrieval for Sentinel-2 is assessed relative to Landsat-9.

III. Methods

The generic workflow of this study is summarized in Fig. 2. The main steps include the following.

1) Preparation of training data by spectral resampling of the hyperspectral R_{rs} data: This step provides the input multispectral R_{rs} data of the DOABLE-Net. The workflow is shown for the OLI data that includes the Pan band allowing for training two different individual networks with and without including this band as an input feature. In the case of MSI, the training does not involve a Pan band. The $R_{\mathrm{rs}}(620)$, created from spectral resampling of the hyperspectral R_{rs}, serves as the response variable of the network.
2) Training the DOABLE-Net with an architecture described in Section III-A.
3) Prediction of the orange band using the trained model and accuracy assessment compared with the reference data. Once the DOABLE-Net is trained, it is applied to the validation OLI or MSI datasets to predict the orange band. The validation set is not seen through the training and allows for assessing the model's generalization capability.

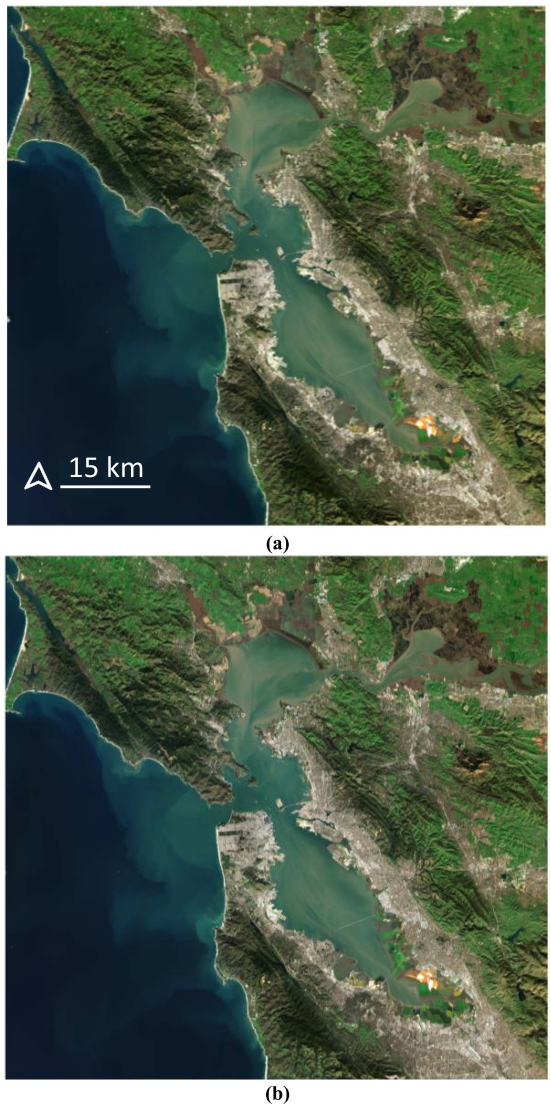

Fig. 1. True color composites of (a) landsat-9 and (b) sentinel-2 images derived from ACOLITE atmospheric correction.

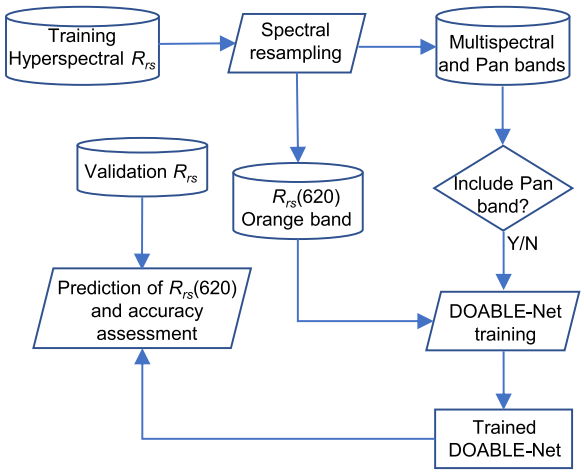

Fig. 2. Workflow of the study based on the proposed DOABLE-Net to retrieve an orange band at 620 nm for OLI and MSI sensors. In the case of MSI, only the version of DOABLE-Net without using a Pan band is applicable.

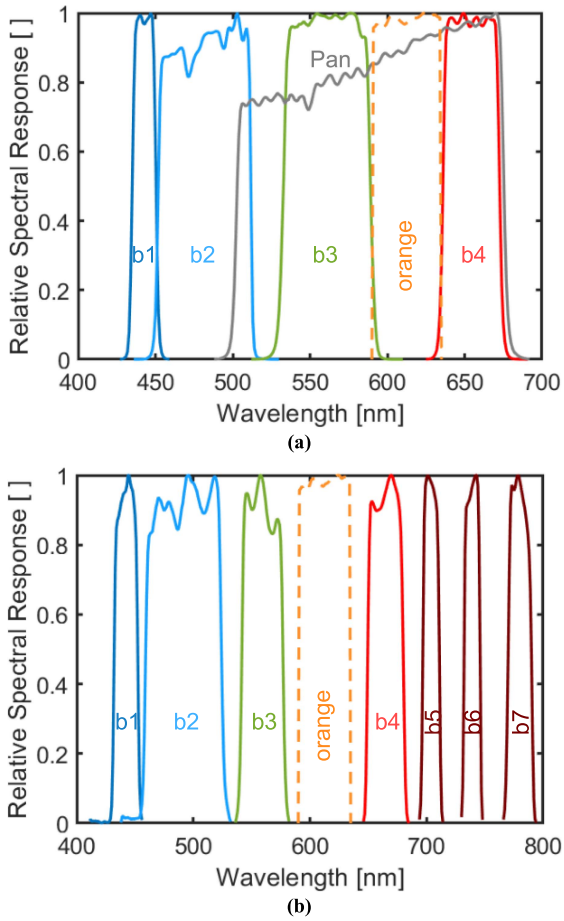

Fig. 3. Relative spectral response of (a) OLI and (b) MSI sensors for the bands up to 800 nm. The dashed orange line shows the virtual orange band (620 nm) taken from the article presented in [9].

The rest of the section describes the details of each step. We first elaborate the proposed DOABLE-Net method. Then, the Pan-based method is presented in its original and calibrated versions. Finally, the accuracy metrics are provided for assessing results and consistency analysis among the methods.

A. *Proposed DOABLE-Net*

Neural networks (NNs) have been proven to be powerful in learning complex and nonlinear relations between input features and the target parameter [31]. Moreover, deep models handle the feature extraction automatically and can leverage low- and high-level features [32], [33]. Given the nonlinear relationships of R_{rs} at different wavelengths [19], we also build upon a deep NN architecture to enable predicting the orange band from multispectral R_{rs} data.

An initial step for the development of supervised deep models is the preparation of the training dataset. The hyperspectral training data (see Section II) are resampled with the Landsat-8/9 and Sentinel-2 spectral response functions for the bands up to 800 nm. Every multispectral band is calculated by a weighted averaging of the hyperspectral R_{rs} data. The weights are taken from the spectral response function of either OLI or MSI sensors. Similarly, an orange band is obtained from the hyperspectral

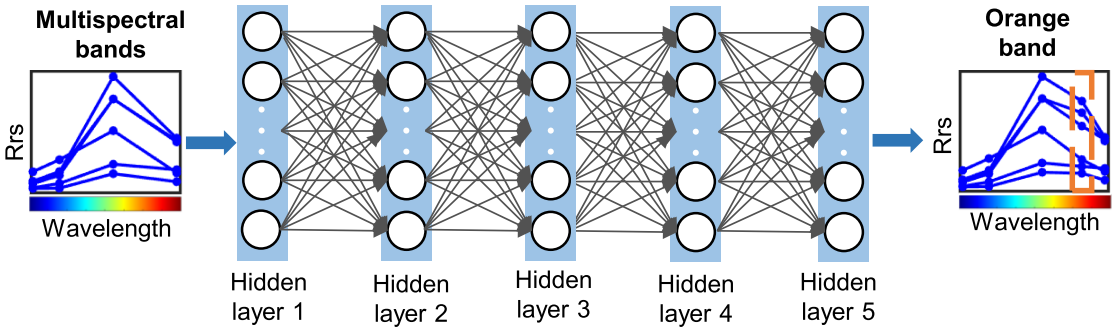

Fig. 4. Architecture of the proposed DOABLE-Net for predicting the orange band, $R_{rs}(620)$, using multispectral bands of landsat-8/9 or sentinel-2.

data equivalent to the one considered in the Pan-based method [9]. The relative spectral responses of OLI and MSI sensors and the orange band centered at 620 nm are shown in Fig. 3 [34].

The proposed DOABLE-Net comprises multiple fully connected hidden layers. We adapted this architecture from the article presented in [19] that provided promising results in a similar task, i.e., retrieving near-blue ultraviolet R_{rs} from visible bands. Although the types of layers are taken from the article presented in [19], the number of layers and hidden units are optimized through a tuning process. Similar architectures have shown outstanding results also in other aquatic applications (e.g., bathymetry and water quality retrieval), dealing with comparable regression tasks using data without spatial context information [35], [36], [37], [38].

The first layer is supplied with the multispectral R_{rs} bands of either Landsat-8/9 or Sentinel-2 created from the hyperspectral training data. Thus, the input layer involves four R_{rs} bands of Landsat-8/9 [b1–b4, Fig. 3(a)] or seven bands of Sentinel-2 [b1–b7, Fig. 3(b)]. We have implemented another version of DOABLE-Net for Landsat-8/9 that includes the Pan band as an extra input feature (five features in total). The latter version is termed DOABLE-Net-Pan, hereafter, which is applicable only to Landsat-8/9 data. The last fully connected layer is connected to the output layer that produces the $R_{rs}(620)$ as the response parameter of the network. As typical fully connected architectures, every layer multiplies the input by a weight matrix followed by adding a bias vector. The working mechanism of such NNs is well-documented in the literature [39].

The hyperparameters of DOABLE-Net, including the number of layers, the number of neurons in each layer, and the type of activation function, are tuned through a Bayesian optimization [40], [41]. The hyperparameter tuning investigates a search space for every parameter to find the optimal value that minimizes the cross-validation error. The tuned architecture of the proposed DOABLE-Net for Landsat-8/9 consists of five fully connected layers with 250, 174, 96, 85, and 22 neurons at each layer, successively.

A similar architecture is defined for Sentinel-2 with layer sizes of 239, 205, 105, 34, and 15 neurons. The tuning is performed in a search space with layer sizes of up to 500 neurons that are large enough for this task [19]. The rectified linear unit is considered as an activation function that effectively tackles gradient explosion and disappearance [42]. A schematic representation of the DOABLE-Net architecture is shown in Fig. 4.

TABLE II
ACCURACY OF ORANGE BAND RETRIEVAL FROM IN SITU VALIDATION DATASETS BASED ON DOABLE-NET AND PAN-BASED METHODS

Sensor	Method	Metrics	SpecWA	SeaSWIR	NORCOHAB	Belgian-Dutch Lakes
Landsat-8/9	DOABLE-Net	R^2	0.98	1	0.97	0.98
		RMSE [sr^{-1}]	7.1×10^{-4}	5.9×10^{-4}	5.6×10^{-4}	7.2×10^{-4}
		NRMSE [%]	2.6	3.8	6.7	2.5
		Bias	0.98	0.95	1.17	1.01
	DOABLE-Net-Pan	R^2	0.98	1	0.97	0.98
		RMSE [sr^{-1}]	6.4×10^{-4}	5.7×10^{-4}	5×10^{-4}	6.8×10^{-4}
		NRMSE [%]	2.3	3.7	6.1	2.3
		Bias	1	1.05	1.13	1.02
	Pan-based	R^2	0.97	1	0.96	0.99
		RMSE [sr^{-1}]	7.5×10^{-4}	5.8×10^{-4}	7.6×10^{-4}	5×10^{-4}
		NRMSE [%]	2.9	3.8	9.2	2
		Bias	0.96	0.96	1.23	0.96
	Pan-based-Cal	R^2	0.96	1	0.96	0.98
		RMSE [sr^{-1}]	7.3×10^{-4}	10×10^{-4}	12×10^{-4}	6×10^{-4}
		NRMSE [%]	2.8	6.5	14.5	2.1
		Bias	1.03	1.10	1.41	1.02
Sentinel-2	DOABLE-Net	R^2	0.97	1	0.96	0.96
		RMSE [sr^{-1}]	8.1×10^{-4}	3×10^{-4}	6.9×10^{-4}	8.2×10^{-4}
		NRMSE [%]	3	1.9	8.3	2.8
		Bias	1.01	0.99	1.21	1.01

B. Original and Calibrated Pan-Based Method for Retrieving Orange Band for Landsat-8/9

The Pan-based method is the only study in the literature that addresses the orange band retrieval for Landsat-8/9 [9]. This approach relies on a simple linear regression model considering the green (R_{rs}^{Green}), red (R_{rs}^{Red}), and Pan (R_{rs}^{Pan}) bands as predictors of the orange band $R_{rs}(620)$ formulated in (1). The linear model is originally trained using 428 in situ hyperspectral R_{rs} spectra resampled to Landsat-8 bands and the orange band. The training set represents a wide range of bio-optical conditions in Belgian

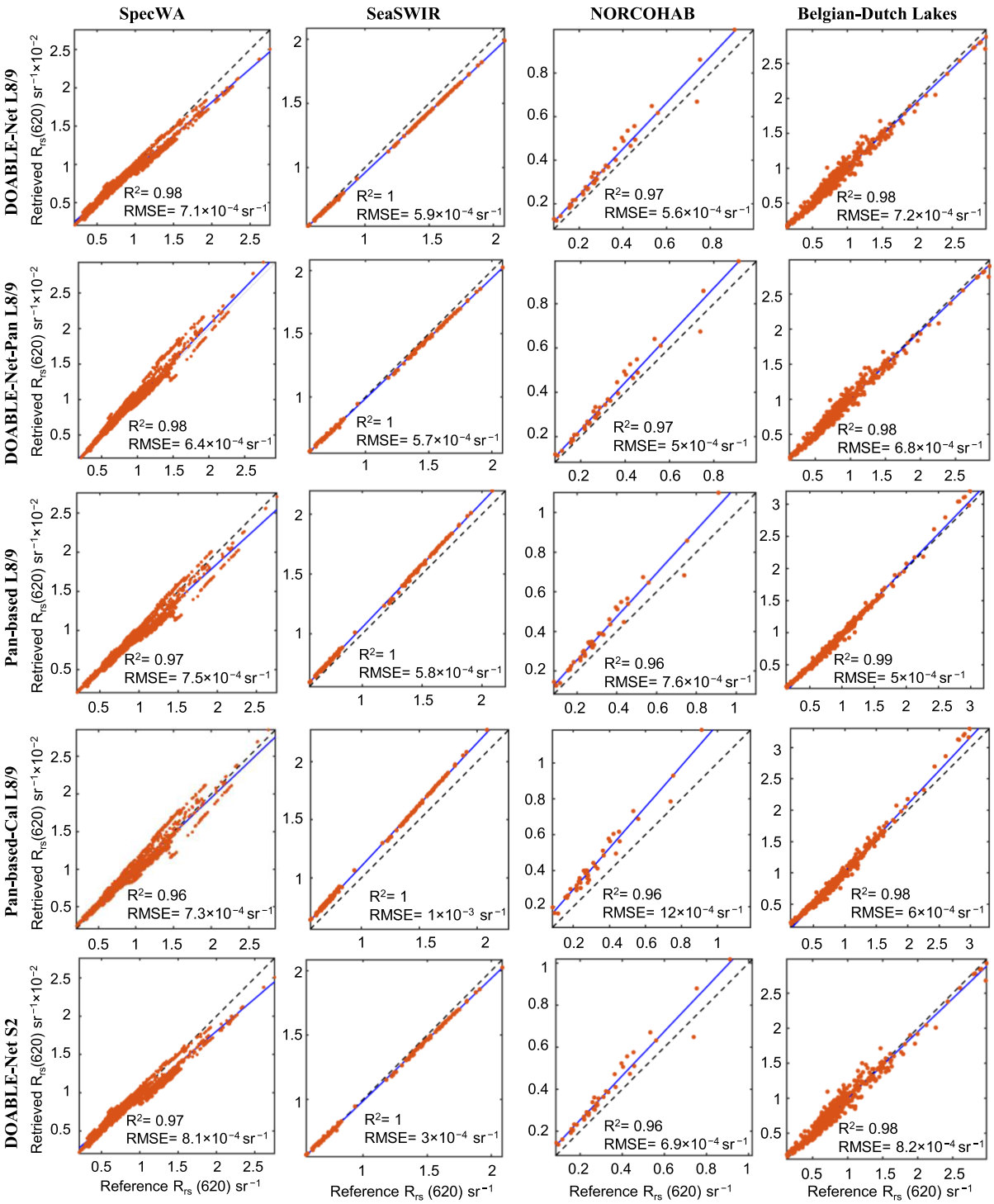

Fig. 5. Validation of the proposed DOABLE-Net in retrieving $R_{\mathrm{rs}}(620)$ compared with the pan-based method for landsat-8/9 (L8/9) and Sentinel-2 (S2) in situ data measured at diverse bio-optical conditions. The impact of including the Pan band is investigated on the performance of the proposed method (DOABLE-Net-Pan).

and Dutch lakes [9], which is used as a validation set in this study (see Table I)

$$R_{\mathrm{rs}}^{\mathrm{Orig}}(620) = 2.2861 \times R_{\mathrm{rs}}^{\mathrm{Pan}} - 0.9467 \times R_{\mathrm{rs}}^{\mathrm{Green}} - 0.1989 \times R_{\mathrm{rs}}^{\mathrm{Red}}. \quad (1)$$

The coefficients of the Pan-based method are assumed to be generic and implemented in ACOLITE processor [34], although one may update the coefficients using other training datasets. In this study, we compare the performance of the Pan-based method with the proposed DOABLE-Net considering both original and calibrated coefficients. The calibration of the Pan-based model is based on the same dataset used for training DOABLE-Net.

Hereafter, we refer to this version of the model as Pan-based-Cal. The calibrated coefficients are provided as follows:

$$R_{\mathrm{rs}}^{\mathrm{Cal}}(620) = 3.0807\times 10^{-4} + 3.703\times R_{\mathrm{rs}}^{\mathrm{Pan}} - 1.6714\times R_{\mathrm{rs}}^{\mathrm{Green}} - 0.82475\times R_{\mathrm{rs}}^{\mathrm{Red}}. \quad (2)$$

Although the Pan-based method is originally developed for Landsat-8, it can be directly applied to Landsat-9 data due to the identical spectral designation of OLI and OLI-2 [43]. However, this method does not apply to Sentinel-2 data that lacks the Pan band. Note that the Pan-based method aims to enhance the spectral resolution and does not involve any pan sharpening to enhance the spatial resolution. Thus, in the image analysis, the Pan band is downsampled to the spatial resolution of multispectral bands (i.e., 30 m).

C. Accuracy Assessment Metrics

We use a set of metrics to quantify the accuracy of orange band retrieval, including coefficient of determination (R^2), root-mean-square error (RMSE), normalized RMSE (NRMSE), and bias

$$R^2 = \frac{\sum_{i=1}^{n}\left(E_i-\bar{O}\right)^2}{\sum_{i=1}^{n}\left(O_i-\bar{O}\right)^2},\ \bar{O} = \frac{1}{n}\sum_{i=1}^{n} O_i \quad (3)$$

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

$$\mathrm{NRMSE}\ [\%] = \frac{\mathrm{RMSE}}{\max(\mathrm{O})-\min(\mathrm{O})}\times 100 \quad (5)$$

$$\mathrm{bias} = 10^{\frac{\sum_{i=1}^{n}\log_{10}(E_i/O_i)}{n}}. \quad (6)$$

E_i and O_i stand for the estimated and observed/known values of the $R_{\mathrm{rs}}(620)$, respectively. The NRMSE indicates the percentage of the RMSE relative to the range of the retrieved parameter (orange band). The bias is computed in a log-transformed space that accounts for the proportionality of the errors [44]. Bias >1 implies an overestimated $R_{\mathrm{rs}}(620)$, whereas bias <1 denotes an underestimation error. For instance, a 15% overestimation in $R_{\mathrm{rs}}(620)$ translates into a bias of 1.15. So, the closer the bias values to 1, the lower the systematic errors. The accuracy metrics are computed for the four validation datasets (see Table I) by comparing the orange band retrievals with the reference in situ observations. In the case of image analysis (San Francisco Bay), the consistency of the maps derived from different methods is assessed by quantifying the pixel-by-pixel agreement in terms of R^2 and normalized root-mean-square difference (NRMSD).

IV. Results and Discussion

The validation scatterplots of retrieved versus measured $R_{\mathrm{rs}}(620)$ are shown in Fig. 5 for the four validation R_{rs} datasets (see Table I). The accuracy statistics based on different metrics are reported in Table II. The proposed DOABLE-Net for Landsat-8/9, without incorporating the Pan band, provides better or comparable retrievals of $R_{\mathrm{rs}}(620)$ relative to the Pan-based method. For instance, DOABLE-Net improved the NRMSE of orange band retrieval by ~3% for the NORCOHAB dataset. Inclusion of the Pan band through DOABLE-Net-Pan improves the results but not significantly. In the case of the Belgian-Dutch lakes dataset, the DOABLE-Net provides very close $R_{\mathrm{rs}}(620)$ retrievals compared with the Pan-based method, even without using the Pan band. This is a promising result for the DOABLE-Net as this dataset was originally used for developing the Pan-based method, whereas it is totally independent for DOABLE-Net (unseen through training). The results also indicate that the calibration of the Pan-based method slightly drops the accuracy of the $R_{\mathrm{rs}}(620)$ retrieval. For instance, the NRMSE for SeaSWIR dataset increases from 3.8% to 6.5% when using the calibrated version of the Pan-based method instead of the original one (see

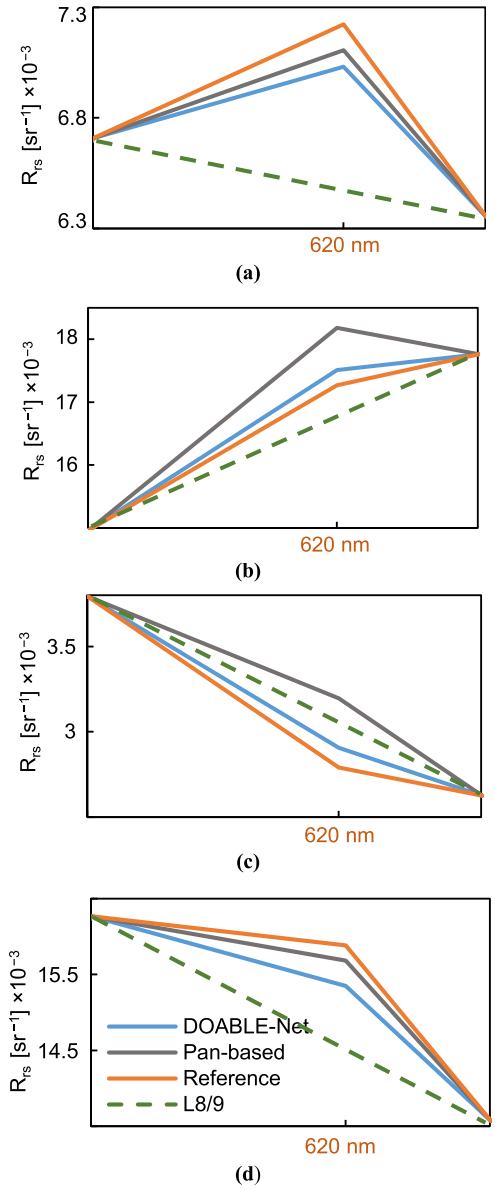

Fig. 6. Samples of retrieved Landsat-8/9 spectra around 620 nm for (a) SpecWA, (b) SeaSWIR, (c) NORCOHAB, and (d) Belgian-Dutch lakes datasets compared with the reference spectra.

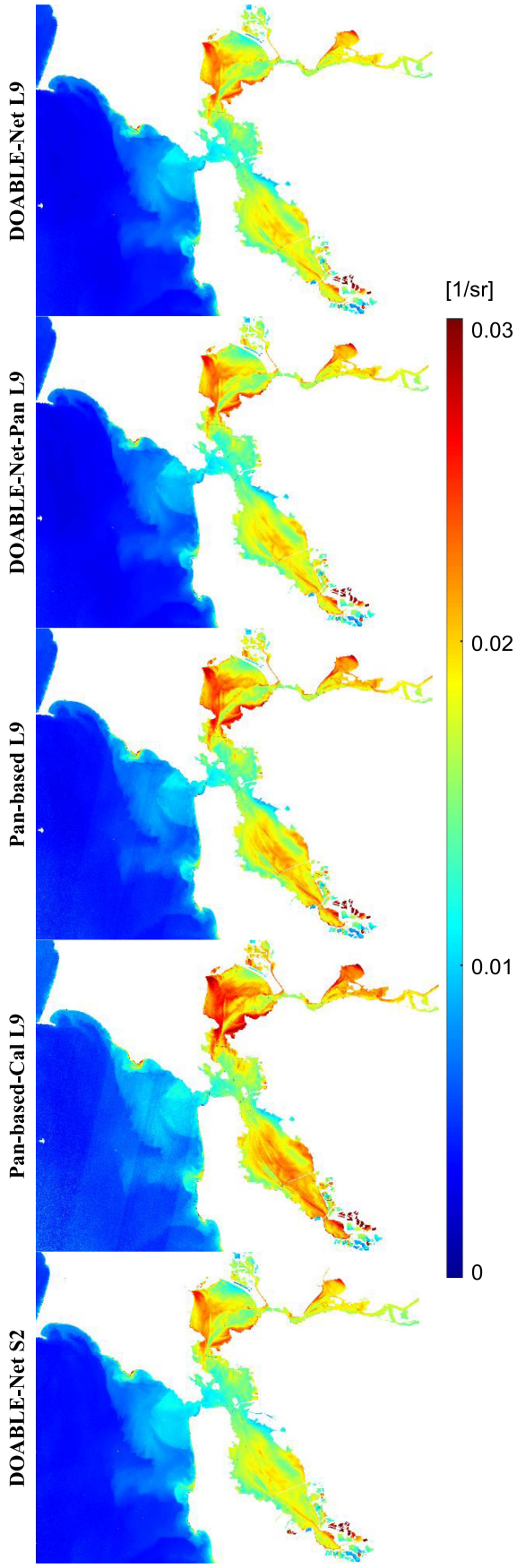

Fig. 7. $R_{rs}(620)$ band retrieved from landsat-9 (L9) and sentinel-2 (S2) imagery of San Francisco Bay based on DOABLE-Net and pan-based methods.

Table II). This can be attributed to the nonlinear and complex interrelationships among the R_{rs} data at different wavelengths that are more pronounced in the large training set from diverse biooptical conditions. So, deep models, such as the proposed DOABLE-Net, are preferred to provide robust retrievals. The orange band retrievals from the Sentinel-2 data are comparable with those of Landsat-8/9 data (see Fig. 5 and Table II). This finding indicates that the proposed DOABLE-Net can reasonably retrieve the $R_{rs}(620)$ for Sentinel-2 data in the lack of a Pan band.

Examples of randomly selected Landsat-8/9 spectra from different datasets, including the retrieved virtual orange band, are shown in Fig. 6 compared with the reference spectra from the hyperspectral in situ data. The original Landsat-8/9 spectra, without the orange band, are also illustrated to better understand the impact of the retrieved $R_{rs}(620)$ in capturing the spectral shape of water-leaving radiance. As evident, the proposed DOABLE-Net, even without incorporating the Pan band, provides stronger or comparable agreements with the reference spectra relative to the Pan-based method. As evident, the retrieved $R_{rs}(620)$ based on DOABLE-Net enables better capturing of the spectral shape around 620 nm.

Fig. 7 shows the $R_{rs}(620)$ maps retrieved from the Landsat-9 and Sentinel-2 imagery acquired over San Francisco Bay. There are strong agreements among different cases derived from DOABLE-Net and Pan-based methods ($R^2 > 0.97$, RMSD < 0.0018 sr^{-1}, NRMSD $< 6\%$, and $0.98 <$bias< 1.03 for any pair of maps). The orange band retrieved based on the Pan-based-Cal for the Landsat-9 image is slightly noisy in the coastal area implying again that the calibration of the linear regression model based on large training data (2) is not helpful. This experiment also conveys that the inclusion of the Pan band has minimal impacts on retrieving the $R_{rs}(620)$ from the Landsat-9 image over diverse water types. Similarly, the orange band retrieval from Sentinel-2 in the lack of a Pan band is comparable with that of Landsat-9.

V. Conclusion

Multispectral satellite sensors, such as OLI and MSI, are yet the main sources of imagery for studying inland waters from space. However, their limited spectral bands hinder the retrieval of specific bio-optical attributes. The lack of an orange band is a key limitation in these sensors' spectral designation, making the detection and quantification of cyano-HABs challenging. This issue has motivated the idea of retrieving a virtual $R_{rs}(620)$ with some potentials previously demonstrated based on the Pan-based method [9]. However, this method is not applicable in the lack of a Pan band, which is the case for Sentinel-2. Furthermore, the Pan-based method is based on a simple linear model, neglecting the complexity and nonlinearity of interrelationships among the R_{rs} data at different portions of the spectrum. We proposed DOABLE-Net, which is based on a deep regression model that allows for automatically capturing the informative features from multispectral data to predict the $R_{rs}(620)$ either with or without using a Pan band. Although the results indicate the minimal impact of the Pan band on the performance of the proposed DOABLE-Net, there is no reason to exclude this band when it is

available, such as for Landsat-8/9 data. However, these results encourage retrieving the orange band for Sentinel-2 without a Pan band. The calibration of the Pan-based model using the same training set of DOABLE-Net degraded the performance of $R_{rs}(620)$ retrieval. This is because our training set is much larger ($\sim$ 250 times) and more diverse than the dataset used in computing the Pan-based method's original coefficients, amplifying the nonlinearity and complexity of the orange band retrieval. Thus, the proposed DOABLE-Net provides a more robust means of retrieving $R_{rs}(620)$ by automatically extracting nonlinear features.

We focused on developing a methodology for $R_{rs}(620)$ retrieval that is evaluated with diverse datasets. Future studies need to assess the impact of the proposed DOABLE-Net in retrieving phycocyanin concentration and mapping cyano-HABs. However, the benefits of $R_{rs}(620)$ retrievals from the Pan-based method have previously been demonstrated in some studies [16], [17]. Thus, such benefits can also be expected with more promises for DOABLE-Net, given its robustness and less dependency on the Pan band. The DOABLE-Net can also potentially be extended for any other multispectral sensor lacking the orange band that requires more investigation. Although we employed a five-layer fully connected network, other deep architectures can also be assessed through orange band retrieval.

REFERENCES

[1] M. Niroumand-Jadidi, F. Bovolo, M. Bresciani, P. Gege, and C. Giardino, "Water quality retrieval from landsat-9 (OLI-2) imagery and comparison to sentinel-2," *Remote Sens.*, vol. 14, no. 18, 2022, Art. no. 4596.

[2] I. Caballero, M. Roca, J. Santos-Echeandía, P. Bernárdez, and G. Navarro, "Use of the sentinel-2 and landsat-8 satellites for water quality monitoring: An early warning tool in the mar menor coastal lagoon," *Remote Sens.*, vol. 14, no. 12, 2022, Art. no. 2744.

[3] C. H. Hansen, S. J. Burian, P. E. Dennison, and G. R. Williams, "Spatiotemporal variability of lake water quality in the context of remote sensing models," *Remote Sens.*, vol. 9, no. 5, Apr. 2017, Art. no. 409.

[4] M. Niroumand-Jadidi, F. Bovolo, L. Bruzzone, and P. Gege, "Intercomparison of methods for chlorophyll-a retrieval: Sentinel-2 time-series analysis in Italian lakes," *Remote Sens.*, vol. 13, no. 12, Jun. 2021, Art. no. 2381.

[5] M. Xu et al., "Implementation strategy and spatiotemporal extensibility of multipredictor ensemble model for water quality parameter retrieval with multispectral remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4200616.

[6] P. D. Hunter, A. N. Tyler, L. Carvalho, G. A. Codd, and S. C. Maberly, "Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2705–2718, 2010.

[7] S. Mishra, R. P. Stumpf, B. A. Schaeffer, P. J. Werdell, K. A. Loftin, and A. Meredith, "Measurement of cyanobacterial bloom magnitude using satellite remote sensing," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 18310.

[8] J. L. Wolny et al., "Current and future remote sensing of harmful algal blooms in the chesapeake bay to support the shellfish industry," *Front. Mar. Sci.*, vol. 7, 2020, Art. no. 337.

[9] A. Castagna, S. Simis, H. Dierssen, Q. Vanhellemont, K. Sabbe, and W. Vyverman, "Extending landsat 8: Retrieval of an orange contra-band for inland water quality applications," *Remote Sens.*, vol. 12, no. 4, 2020, Art. no. 637.

[10] I. Ogashawara, "The use of sentinel-3 imagery to monitor cyanobacterial blooms," *Environments*, vol. 6, no. 6, 2019, Art. no. 60.

[11] M. Bresciani et al., "Application of new hyperspectral sensors in the remote sensing of aquatic ecosystem health: Exploiting PRISMA and DESIS for four Italian lakes," *Resources*, vol. 11, no. 2, 2022, Art. no. 8.

[12] M. Niroumand-Jadidi, F. Bovolo, and L. Bruzzone, "Water quality retrieval from PRISMA hyperspectral images: First experience in a turbid lake and comparison with sentinel-2," *Remote Sens.*, vol. 12, no. 23, Dec. 2020, Art. no. 3984.

[13] T. Kutser, "Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters," *Int. J. Remote Sens.*, vol. 30, no. 17, pp. 4401–4425, 2009.

[14] S. L. Sharp, A. L. Forrest, K. Bouma-Gregson, Y. Jin, A. Cortés, and S. G. Schladow, "Quantifying scales of spatial variability of cyanobacteria in a large, Eutrophic lake using multiplatform remote sensing tools," *Front. Environ. Sci.*, vol. 9, 2021, Art. no. 612934.

[15] Z.-P. Lee, M. Darecki, K. L. Carder, C. O. Davis, D. Stramski, and W. J. Rhea, "Diffuse attenuation coefficient of downwelling irradiance: An evaluation of remote sensing methods," *J. Geophys. Res. Oceans*, vol. 110, no. C2, 2005, Art. no. C02017.

[16] A. Kumar, D. R. Mishra, and N. Ilango, "Landsat 8 virtual orange band for mapping cyanobacterial blooms," *Remote Sens.*, vol. 12, no. 5, 2020, Art. no. 868.

[17] I. Ogashawara, L. Li, C. Howard, and G. K. Druschel, "Monitoring phycocyanin with landsat 8/operational land imager orange contra-band," *Environments*, vol. 9, no. 3, 2022, Art. no. 40.

[18] Q. Vanhellemont, "Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the landsat and sentinel-2 archives," *Remote Sens. Environ.*, vol. 225, pp. 175–192, 2019.

[19] Y. Wang, Z. Lee, J. Wei, S. Shang, M. Wang, and W. Lai, "Extending satellite ocean color remote sensing to the near-blue ultraviolet bands," *Remote Sens. Environ.*, vol. 253, 2021, Art. no. 112228.

[20] Z. Lee, S. Shang, C. Hu, and G. Zibordi, "Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors," *Appl. Opt.*, vol. 53, no. 15, pp. 3301–3310, May 2014.

[21] Z. Li et al., "A reconstruction method for hyperspectral remote sensing reflectance in the visible domain and applications," *J. Geophys. Res. Oceans*, vol. 123, no. 6, pp. 4092–4109, 2018.

[22] G. Yulong et al., "Hyperspectral reconstruction method for optically complex inland waters based on bio-optical model and sparse representing," *Remote Sens. Environ.*, vol. 276, 2022, Art. no. 113045.

[23] C. D. Mobley, *Light and Water: Radiative Transfer in Natural Waters*. New York, NY, USA: Academic, 1994.

[24] P. Gege, "The water color simulator WASI: An integrating software tool for analysis and simulation of optical in situ spectra," *Comput. Geosci.*, vol. 30, no. 5, pp. 523–532, Jun. 2004.

[25] G. H. Tilstone et al., "Consistency between satellite ocean colour products under high coloured dissolved organic matter absorption in the Baltic Sea," *Remote Sens.*, vol. 14, no. 1, 2022, Art. no. 89.

[26] P. M. Maier and S. Keller, "SpecWa: Spectral remote sensing data and chlorophyll a values of inland waters," GFZ Data Services, Potsdam, Germany, 2020.

[27] E. Knaeps et al., "The SeaSWIR dataset," *Earth Syst. Sci. Data*, vol. 10, pp. 1439–1449, 2018.

[28] S. P. Garaba, M. R. Wernand, and O. Zielinski, "Downward irradiance during the North Sea Coast harmful algal bloom (NORCOHAB II) HEINCKE cruise HE302," Pangaea, 2011.

[29] A. Castagna et al., "Dataset of optical and biogeochemical properties of diverse Belgian inland and coastal waters," Pangaea, 2022.

[30] Q. Vanhellemont, "Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre- and decametre-scale satellite imagery using autonomous hyperspectral radiometry," *Opt. Express*, vol. 28, no. 20, pp. 29948–29965, Sep. 2020.

[31] F. Murtagh, "Multilayer perceptrons for classification and regression," *Neurocomputing*, vol. 2, no. 5/6, pp. 183–197, Jul. 1991.

[32] M. Gevrey, I. Dimopoulos, and S. Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models," *Ecol. Model.*, vol. 160, no. 3, pp. 249–264, 2003.

[33] H. Liu et al., "Estimating ultraviolet reflectance from visible bands in ocean colour remote sensing," *Remote Sens. Environ.*, vol. 258, 2021, Art. no. 112404.

[34] Q. Vanhellemont, "Acolite/data/RSR at main acolite/acolite Github," Accessed on: May 9, 2022. [Online]. Available: https://github.com/acolite/acolite/tree/main/data/RSR

[35] M. Niroumand-Jadidi, C. J. Legleiter, and F. Bovolo, "Bathymetry retrieval from CubeSat image sequences with short time lags," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 112, 2022, Art. no. 102958.

[36] M. Niroumand-Jadidi and F. Bovolo, "Temporally transferable machine learning model for total suspended matter retrieval from sentinel-2," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. V-3–2022, pp. 339–345, 2022.
[37] M. Niroumand-Jadidi, C. J. Legleiter, and F. Bovolo, "River bathymetry retrieval from landsat-9 images based on neural networks and comparison to superdove and sentinel-2," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 5250–5260, 2022.
[38] J. Chen et al., "Remote sensing estimation of chlorophyll-A in case-II waters of coastal areas: Three-band model versus genetic algorithm–artificial neural networks model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3640–3658, 2021.
[39] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, Mar. 1996.
[40] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.* 2012, vol. 25.
[41] M. A. Gelbart, J. Snoek, and R. P. Adams, "Bayesian optimization with unknown constraints," Mar. 2014, *arXiv:1403.5607*.
[42] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," *CoRR*, 2015, *arXiv:1502.01852*.
[43] J. G. Masek et al., "Landsat 9: Empowering open science and applications through continuity," *Remote Sens. Environ.*, vol. 248, Oct. 2020, Art. no. 111968.
[44] B. N. Seegers, R. P. Stumpf, B. A. Schaeffer, K. A. Loftin, and P. J. Werdell, "Performance metrics for the assessment of satellite data products: An ocean color case study," *Opt. Express*, vol. 26, no. 6, pp. 7404–7422, Mar. 2018.

Milad Niroumand-Jadidi (Member, IEEE) received the B.Sc. degree in geomatics engineering from the University of Tabriz, Tabriz, Iran, in 2009, and the M.Sc. degree in remote sensing engineering from the K.N. Toosi University of Technology, Tehran, Iran, in 2013. He received the joint Ph.D. degree in civil and environmental engineering from the University of Trento, Trento, Italy, and Freie Universität Berlin, Berlin, Germany.

Since 2017, he has been a Postdoctoral Researcher with the Remote Sensing for Digital Earth Unit of the Digital Society Center, Fondazione Bruno Kessler, Trento, Italy. He has authored or coauthored more than 15 peer-reviewed articles in top journals and presented numerous papers at international conferences. His main research interests include the development of methods and applications for remote sensing of inland and coastal waters from optical data. He focuses on both physics-based and machine learning models to retrieve information on water quality and bathymetry.

Dr. Niroumand-Jadidi was a recipient of several international awards. Among those, he has been awarded a DLR-DAAD fellowship to conduct research with the German Aerospace Center (DLR) as a Visiting Scientist (May−August 2022). He recieved a Travel Grant to attend Ocean Optics XXV Conference (Vietnam, 2022) from European Operational Satellite Agency for monitoring weather, climate, and the environment from space. He was a recipient of the Best Young Author Award from the International Society for Photogrammetry and Remote Sensing in 2021, the Best Paper Award from SPIE Remote Sensing Conferences for three years in a row (2015 in Toulouse, 2016 in Edinburgh, and 2017 in Warsaw), the award for potential long-range contribution to the field of optics and photonics from SPIE (The International Society for Optics and Photonics) in 2016, and Alexander Goetz Instrument Support Award 2016, which gave him access to a field spectroradiometer during his Ph.D. research. He serves on the editorial board for the journal *Remote Sensing* and has also been a special issue Editor. He is a referee for several international journals, including IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Remote Sensing of Environment*, *Remote Sensing*, *Journal of Hydrology*, and *International Journal of Remote Sensing*.

Francesca Bovolo (Senior Member, IEEE) received the B.S. (Laurea) and M.S. (Laurea Specialistica) degrees (*summa cum laude*) in telecommunication engineering, and the Ph.D. degree in communication and information technologies from the University of Trento, Trento, Italy, in 2001, 2003, and 2006, respectively.

She was a Research Fellow with the University of Trento, until 2013. She is currently the Founder and the Head of Remote Sensing for Digital Earth Unit, Fondazione Bruno Kessler, Trento, Italy, and a member of Remote Sensing Laboratory, Trento, Italy. She is one of the Co-Investigators of the Radar for Icy Moon Exploration Instrument of the European Space Agency Jupiter Icy Moons Explorer and member of the science study team of the EnVision mission to Venus. She conducts research on these topics within the context of several national and international projects. Her research interests include remote-sensing image processing, multitemporal remote sensing image analysis, change detection in multispectral, hyperspectral, and synthetic aperture radar images and very high-resolution images, time-series analysis, content-based time-series retrieval, domain adaptation, and light detection and ranging and radar sounders.

Dr. Bovolo is a member of the program and scientific committee of several international conferences and workshops. She was the recipient of the First Place in the Student Prize Paper Competition of the 2006 IEEE International Geoscience and Remote Sensing Symposium (Denver, 2006). She was the Technical Chair of the Sixth International Workshop on the Analysis of Multitemporal Remote-Sensing Images (MultiTemp 2011 and 2019). She has been a Co-Chair of the SPIE International Conference on Signal and Image Processing for Remote Sensing since 2014. She is the Publication Chair for the International Geoscience and Remote Sensing Symposium in 2015. She has been an Associate Editor for the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* since 2011 and the Guest Editor for the Special Issue on Analysis of Multitemporal Remote Sensing Data of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. She is a referee for several international journals.