

# A revision of NASA SeaDAS atmospheric correction algorithm over turbid waters with artificial Neural Networks estimated remote-sensing reflectance in the near-infrared

Junwei Wang<sup>a</sup>, Yongchao Wang<sup>a</sup>, Zhongping Lee<sup>a,b,\*</sup>, Daosheng Wang<sup>a</sup>, Shuguo Chen<sup>c</sup>, Wendian Lai<sup>a</sup>

<sup>a</sup> State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China

<sup>b</sup> School for the Environment, University of Massachusetts Boston, Boston, MA 02125, USA

<sup>c</sup> College of Marine Technology, Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China

## ARTICLE INFO

### Keywords:

Ocean color remote sensing  
Atmospheric correction  
Artificial Neural Network  
Near-infrared  
Coastal turbid waters

## ABSTRACT

For atmospheric correction over turbid waters, due to non-negligible water-leaving radiance ( $L_w$ ) in the near-infrared (NIR), measurements in the short-wave infrared (SWIR) are usually required to achieve reliable remote-sensing reflectance ( $R_{rs}$ ). But several ocean color satellite sensors, such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and other small satellites, have no bands in the SWIR domain. We here present an atmospheric correction algorithm (termed as ACA<sub>NIR-NN</sub>) based on NASA SeaDAS (version 7.5.3), which can achieve atmospheric correction seamlessly over clear and turbid waters, even for sensors having no spectral bands in SWIR. Specifically, ACA<sub>NIR-NN</sub> uses estimated  $R_{rs}$ (NIR) from available  $R_{rs}$  in the visible bands with a specifically designed artificial Neural Networks to carry out atmospheric correction, and the performance of ACA<sub>NIR-NN</sub> is evaluated over eight coastal locations having ground measurements by the Aerosol Robotic Network-Ocean Color (AERONET-OC) system. It is found that the Mean Absolute Percent Difference (MAPD) of  $R_{rs}$  retrievals by ACA<sub>NIR-NN</sub> for this dataset is smaller by a factor of two or more than that by the standard SeaDAS algorithm (termed as ACA<sub>NIR-bio</sub>) for each band, especially for  $R_{rs}(412)$  and  $R_{rs}(443)$ , which is 7.5% and 7.7%, respectively, from ACA<sub>NIR-NN</sub>, but they are 44.0% and 27.5% from ACA<sub>NIR-bio</sub>. We further demonstrated the applicability of ACA<sub>NIR-NN</sub> to SeaWiFS measurements over turbid waters, where consistent  $R_{rs}$  products were also obtained compared to that generated from the same-day MODerate resolution Imaging Spectrometer (MODIS) measurements using SWIR bands. These results indicate that ACA<sub>NIR-NN</sub> can generate reliable  $R_{rs}$  over turbid coastal areas, as well as clear ocean waters, for sensors having no SWIR bands.

## 1. Introduction

Due to the coupling role of climate change and anthropogenic activities on the ecological environments, it is becoming increasingly crucial to wisely manage and monitor coastal water environments (Min et al. 2020; Murray et al. 2018; Pettorelli et al. 2018; Zou et al. 2011), a task that requires adequate and efficient observations of these vital ecosystems. With broad coverage and frequent observations, satellite remote sensing is an indispensable system to meet this data demand (Mouw et al. 2015). For ocean color remote sensing, while in principle it is possible to derive the properties of the atmosphere and ocean

simultaneously from satellite measurements (Chomko et al. 2003; Kuchinke et al. 2009; Steinmetz et al. 2011; Wang et al. 2021a), or directly to estimate water properties with artificial Neural Networks (NNs) (Fan et al. 2017; Fan et al. 2021; Schroeder et al. 2007), the commonly adopted strategy by the community is a two-steps scheme, which removes the contributions from the atmosphere before deriving water's optical and biogeochemical properties with various inversion algorithms from the remote-sensing reflectance ( $R_{rs}$ ) obtained from the first step (Hu et al. 2012; Shang et al. 2019; Werdell et al. 2013; Yu et al. 2019). As such, the performance of the first step, termed as atmospheric correction (AC), is critical to obtaining reliable products related to water

\* Corresponding author at: School for the Environment, University of Massachusetts Boston, Boston, MA 02125, USA.

E-mail addresses: [junweiwang@stu.xmu.edu.cn](mailto:junweiwang@stu.xmu.edu.cn) (J. Wang), [yongchaowang@stu.xmu.edu.cn](mailto:yongchaowang@stu.xmu.edu.cn) (Y. Wang), [ZhongPing.Lee@umb.edu](mailto:ZhongPing.Lee@umb.edu) (Z. Lee), [dswang@xmu.edu.cn](mailto:dswang@xmu.edu.cn) (D. Wang), [chenshuguo@ouc.edu.cn](mailto:chenshuguo@ouc.edu.cn) (S. Chen), [wendian\\_lai@163.com](mailto:wendian_lai@163.com) (W. Lai).

<https://doi.org/10.1016/j.isprsjprs.2022.10.014>

Received 10 February 2022; Received in revised form 11 October 2022; Accepted 21 October 2022

Available online 7 November 2022

0924-2716/© 2022 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

quality (Bailey et al. 2010; Gordon and Wang 1994; Ruddick et al. 2000).

Remote-sensing reflectance is defined as the ratio of water-leaving radiance ( $L_w$ ) to downwelling irradiance just above the surface ( $E_d(0^+)$ ), which is equivalent to the ratio of normalized water-leaving radiance ( $L_{wn}$ ) to extraterrestrial solar irradiance ( $F_0$ ) (Thuillier et al. 1998).  $L_w$  or  $L_{wn}$  is retrieved from the radiance ( $L_t$ ) measured at the top-of-atmosphere (TOA) after eliminating the contributions from air molecules and aerosols. For open ocean waters,  $L_w$  in the near-infrared (NIR) bands can be considered negligible due to the high absorption by pure (sea)water in this spectral range and low concentrations of suspended particulate matters, which is referred to as the black pixel (BP) assumption (Gordon and Wang 1994). Therefore, the aerosol type and contributions could be estimated using  $L_t$  data in the NIR bands along with pre-calculated look-up tables (LUTs). However, it has difficulties in turbid coastal and inland waters where  $L_w$ (NIR) is no longer negligible due to high concentrations of suspended particulate matters (Bailey et al. 2010; IOCCG 2010; Wang and Shi 2007).

To process satellite ocean color observations where  $R_{rs}$ (NIR) cannot be neglected over coastal waters, several alternative schemes have been developed. The commonly used AC algorithm (ACA) is to estimate  $R_{rs}$ (NIR) based on bio-optical models (BOMs) with an iterative scheme (termed as  $ACA_{NIR-bio}$  hereafter) (Bailey et al. 2010; Ibrahim et al. 2019; Stumpf et al. 2002).  $ACA_{NIR-bio}$  has been adopted by the National Aeronautics and Space Administration (NASA) as the default ACA to process measurements from many ocean color sensors, e.g., Sea-viewing Wide Field-of-view Sensor (SeaWiFS), MODerate resolution Imaging Spectrometer (MODIS), and Visible Infrared Imaging Radiometer Suite (VIIRS). However, many studies (Goyens et al. 2013; Jiang and Wang 2014; Ruddick et al. 2000; Shehhi et al. 2017) have shown that such an approach run into difficulties in many turbid waters due to the bio-optical relationships are not universal. To overcome this limitation, taking advantage of the significantly greater absorption coefficient of pure seawater in the shortwave infrared (SWIR) domain, this “BP” assumption was extended to the SWIR bands ( $ACA_{SWIR}$ ) in the recent decade (Pahlevan et al. 2017; Shi and Wang 2007; Vanhellemont 2020; Vanhellemont and Ruddick 2015; Wang 2007; Wang and Shi 2007). With a wide range of coastal measurements, the performance of this method has been validated with other commonly used ACAs for MODIS and SeaWiFS over turbid waters based on the Aerosol Robotic Network-Ocean Color (AERONET-OC) measurements (Carswell et al. 2017; Goyens et al. 2013; Jamet et al. 2011; Pahlevan et al. 2017). However, a key requirement for  $ACA_{SWIR}$  to work is that an ocean color sensor can provide adequate radiance measurements in the SWIR domain. For many satellite ocean color sensors, such as SeaWiFS, Medium Resolution Imaging Spectrometer (MERIS), Ocean and Land Color Instrument (OLCI) onboard Sentinel-3, Geostationary Ocean Color Imager (GOCI), Chinese Ocean Color and Temperature Scanner (COCTS), and the small satellites (e.g., HawkEye, HiSea-2), there are no bands in the SWIR domain, therefore an application of measurements in the NIR bands is still the only option for atmospheric correction over turbid coastal waters for such sensors for the two-steps processing strategy.

In this study, we present a scheme to estimate  $R_{rs}$ (NIR) for atmospheric correction (termed as  $ACA_{NIR-NN}$ ), where  $R_{rs}$ (NIR) of clear to turbid waters is estimated from available  $R_{rs}$  in the visible bands with a specifically designed Neural Networks. In particular, considering the initial  $R_{rs}$ (visible) products are incorrect due to the “BP” assumption no longer valid in turbid coastal waters, this NN system was developed with error-bearing  $R_{rs}$ (visible) to estimate  $R_{rs}$ (NIR), which is termed as NN-EBVR (NN for  $R_{rs}$ (NIR) using error-bearing visible  $R_{rs}$ ). The scheme is evaluated with a wide range of measurements in turbid coastal regions, with its performance also compared with the conventional AC algorithm ( $ACA_{NIR-bio}$  (Bailey et al. 2010)) and the scheme including measurements in the SWIR bands ( $ACA_{NIR-SWIR}$  (Wang and Shi 2007)).

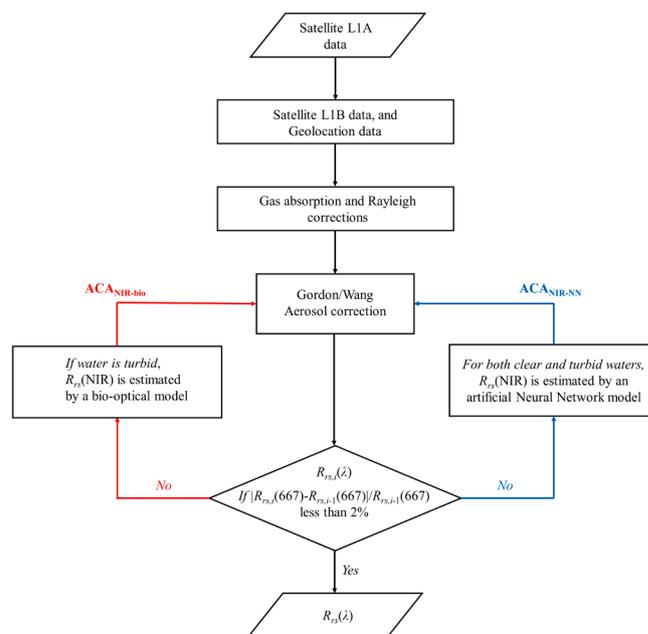


Fig. 1. The overall flowchart of  $ACA_{NIR-bio}$  (left side with red arrows for iteration) and  $ACA_{NIR-NN}$  (right side with blue arrows for iteration). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

## 2. Methods

### 2.1. Background of $ACA_{NIR-bio}$

A satellite sensor measures radiance at the top of atmosphere (TOA), which is commonly converted to reflectance to remove the variations associated with solar radiation, and is defined as,

$$\rho_t = \pi L_t / (\cos(\theta_s) F_0) \quad (1)$$

with  $\rho_t$  the total reflectance at TOA measured by sensors,  $\theta_s$  the solar zenith angle and  $F_0$  the extraterrestrial solar irradiance (Thuillier et al. 1998).

For ocean color remote sensing,  $\rho_t$  is commonly expressed as a sum of the contributions from Rayleigh scattering ( $\rho_r$ ), aerosol scattering ( $\rho_a$ ), the Rayleigh-aerosol inter-scattering ( $\rho_{ra}$ ), sun glint ( $\rho_{sg}$ ) and the remote-sensing reflectance of the water body,

$$\rho_t(\lambda) = t_g(\lambda) [\rho_r(\lambda) + \rho_a(\lambda) + \rho_{ra}(\lambda) + t_s(\lambda) T(\lambda) \rho_{sg} + t_s(\lambda) t_v(\lambda) \pi R_{rs}(\lambda)] \quad (2)$$

Here  $t_g$  is the gas transmittance,  $T$  is the direct transmittance, and  $t_s$  and  $t_v$  are the diffuse atmospheric transmittances for solar radiation reaching the sea surface and water-leaving radiance reaching the sensor, respectively. The influence of whitecaps is ignored here.

For given sun-sensor positions and atmospheric pressure, along with gas information in the atmosphere,  $t_g$  and  $\rho_r$  can be accurately calculated (Gordon et al. 1988; Ibrahim et al. 2018; Wang 2002, 2005), which can then be removed from Eq. (2). The influence of  $\rho_{sg}$  can also be masked or corrected based on Wang and Bailey (2001). A critical step to obtaining accurate  $R_{rs}$  is to remove  $\rho_a$  and  $\rho_{ra}$ , which is commonly accomplished using measurements in the NIR bands (IOCCG 2010), by either assuming  $R_{rs}$ (NIR) as 0 (Gordon and Wang 1994) or estimating  $R_{rs}$ (NIR) through BOMs along with iterative process when the water is too turbid (Bailey et al. 2010; Stumpf et al. 2002). For the ACAs with iterations, it begins by assuming  $R_{rs}$ (NIR) = 0, so that an initial set of  $R_{rs}$ (visible) can be derived and a second set of  $R_{rs}$ (NIR) is estimated based on the concentration of chlorophyll (Chl) and/or suspended particulate matter (SPM) through BOMs (Bailey et al. 2010; Lavender et al. 2005; Xue et al. 2021), with both Chl and SPM derived empirically from this  $R_{rs}$ (visible). The

iteration terminates when  $R_{rs}$  at the red bands changes by less than a threshold, usually set as 2% (Bailey et al. 2010). For most cases, the average number of iterations is 3–4, although 10 is the maximum number of iterations set by the algorithm. The overall workflow of these AC algorithms is shown in Fig. 1.

Among these algorithms, either Chl or SPM is estimated empirically using two  $R_{rs}$  in the visible bands (Bailey et al. 2010; Lavender et al. 2005; Xue et al. 2021), where there are always various levels of uncertainties for different waters, especially for sediment-loaded coastal waters. In addition, the bio-optical models used in the system also contain uncertainties. Thus, a potentially more applicable approach for the estimation of  $R_{rs}(NIR)$  is to (1) use all available  $R_{rs}$  in the visible bands; and (2) by-pass the derivation of Chl and/or SPM as well as the required bio-optical models. We thus, based on SeaDAS (version 7.5.3), propose  $ACA_{NIR-NN}$  to process ocean color satellite measurements that meet the above two desires.

## 2.2. ACA with Neural Networks estimated $R_{rs}(NIR)$

### 2.2.1. The overall strategy of $ACA_{NIR-NN}$

The overall strategy for  $ACA_{NIR-NN}$  is similar to that of  $ACA_{NIR-bio}$ , except that  $R_{rs}(NIR)$  is iteratively estimated from  $R_{rs}(visible)$  by an artificial Neural Networks instead of BOMs. In addition, there is no artificial switch in the data processing by  $ACA_{NIR-NN}$ , as  $R_{rs}(NIR)$  is estimated for each water pixel regardless if the water is clear or turbid. In such a manner, seamless  $R_{rs}$  products can be generated from satellite ocean color data, whereas artificial switch could result in abrupt steps in an image. Fig. 1 illustrates a schematic data flow chart of this system. More importantly, because  $R_{rs}(NIR)$  was assumed as 0 to initiate the calculations, which will then result in inaccurate  $R_{rs}(visible)$  from the first round of calculation, this NN-EBVR was then developed to specifically use inaccurate  $R_{rs}(visible)$  as inputs to estimate  $R_{rs}(NIR)$  as outputs. Using the spectral bands of MODIS Aqua (MODISA) as a demonstration, where the two NIR bands are 748 nm and 869 nm, details of this  $ACA_{NIR-NN}$  are as below:

- (1) In the first round of the AC procedure, where the Gordon and Wang (1994) algorithm is applied, the initial value of  $R_{rs}(869)$  ( $R_{rs}(869)_{.0}$ ) is set to 0, and the initial value of  $R_{rs}(748)_{.0}$  is set in such a manner:

$$R_{rs}(748)_{.0} = (\rho_{rc}(748) - \rho_{rc}(869)) / \pi \quad (3)$$

where  $\rho_{rc}$  represents the Rayleigh-corrected reflectance, which is obtained by SeaDAS (version 7.5.3). In this way, we initially assumed only  $R_{rs}(869)$  as 0, but allowed an estimate for  $R_{rs}(748)_{.0}$  to facilitate the estimation of aerosol properties. With values of  $R_{rs}(748)_{.0}$  and  $R_{rs}(869)_{.0}$  known, the conventional AC procedure is carried out to obtain the first set of  $R_{rs}$  ( $R_{rs-1st}$ ) in the visible domain, which are 412, 443, 488, 531, 547, and 667 nm for the MODIS sensor.

- (2)  $R_{rs-1st}(443, 488, 531, 547, 667)$  are used as the inputs in NN-EBVR to obtain  $R_{rs}(748)_{.2nd}$  and  $R_{rs}(869)_{.2nd}$ .  $R_{rs}(412)$  is excluded in this NN-EBVR system due to that  $R_{rs}(412)$  is mainly driven by colored dissolved organic matter (CDOM), while  $R_{rs}(NIR)$  is driven by suspended sediments. In addition,  $R_{rs}(412)$  of coastal waters could be negative when processed with the conventional AC algorithm.
- (3) These  $R_{rs}(748)_{.2nd}$  and  $R_{rs}(869)_{.2nd}$  values are incorporated into the AC process, yielding a new set of  $R_{rs}(visible)$ .
- (4) These  $R_{rs}(visible)$  subtracted the value of  $R_{rs}(869)_{.2nd}$ , with the results used as the inputs in NN-EBVR to obtain a new set of  $R_{rs}(NIR)$ , i.e.,  $R_{rs}(748)_{.3rd}$  and  $R_{rs}(869)_{.3rd}$ . The reason to subtract  $R_{rs}(869)$  for all  $R_{rs}(visible)$  is because that the training of NN-EBVR is based on inaccurate  $R_{rs}(visible)$  where all  $R_{rs}(869)$  were set as 0.

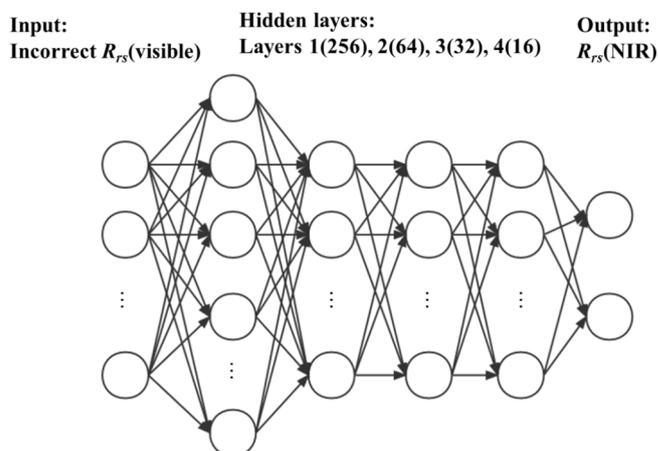


Fig. 2. Deep learning system for the estimation of  $R_{rs}(NIR)$  from incorrect  $R_{rs}(visible)$ : NN-EBVR.

- (5) Steps 3 and 4 are repeated until the retrieved  $R_{rs}(667)$  differ by less than 2% compared to the previous round.

### 2.2.2. NN system for the estimation of $R_{rs}(NIR)$ : NN-EBVR

Similar to all deep-learning systems, the above-mentioned NN-EBVR is composed of one input layer (five inaccurate  $R_{rs}(visible)$  at 443, 488, 531, 547 and 667 nm), various hidden layers that are associated with many numbers of neurons, and one output layer ( $R_{rs}(NIR)$  at 748 and 869 nm in this case). For the NN system, based on the data characteristics, we selected the Keras library (Choi et al. 2017; Ketkar 2017; Moolayil et al. 2019) for the development of NN-EBVR. As a high-level Application Programming Interface for Tensorflow used for deep learning model construction, debugging, assessment and implementation, Keras is an open-source artificial Neural Networks library written in Python (Choi et al. 2017; Ketkar 2017). To determine the optimal number of hidden layers and neurons, we have explored several combination settings. Based on the criteria of performance and computation time, the optimal configuration with hidden layers and neurons has been adopted to train the NN-EBVR. Eventually, this NN-EBVR system is determined to have 4 hidden layers with 256, 64, 32, and 16 neurons (see Fig. 2).

For the training of NN-EBVR, we adopted similar settings (e.g., the activation function, the optimization function and the learning rate) as that in Wang et al., (2021b). Ultimately, while the implementation of the  $ACA_{NIR-NN}$  followed the data flow as in Bailey et al. (2010) that included in SeaDAS (version 7.5.3), this NN-EBVR module replaced the block of estimating  $R_{rs}(NIR)$  that uses Chl and bio-optical models.

## 3. Data

### 3.1. Data for the Neural Networks model

For the development of all Neural Networks or deep-learning-based algorithms, a large and diverse dataset is the key. Here we employ a numerically synthetic dataset for the development of NN-EBVR, which was further assessed by 20% of the synthetic dataset and *in situ* measurements.

An  $R_{rs}(\lambda)$  spectrum is governed by two spectra: the absorption ( $a(\lambda)$ ) and backscattering ( $b_b(\lambda)$ ) coefficients of the water body, which are termed as the inherent optical properties (IOPs) (Preisendorfer 1976). Both  $a(\lambda)$  and  $b_b(\lambda)$  vary significantly from oceanic to coastal environments. Thus, to generate a dataset having a wide range of  $R_{rs}(\lambda)$ , it is key to have a wide range of, and reasonable,  $a(\lambda)$  and  $b_b(\lambda)$ . Both  $a(\lambda)$  and  $b_b(\lambda)$  are composed of contributions from water itself and water constituents, including phytoplankton pigments, CDOM and detritus-minerals. As most details for this synthesizing process are available in

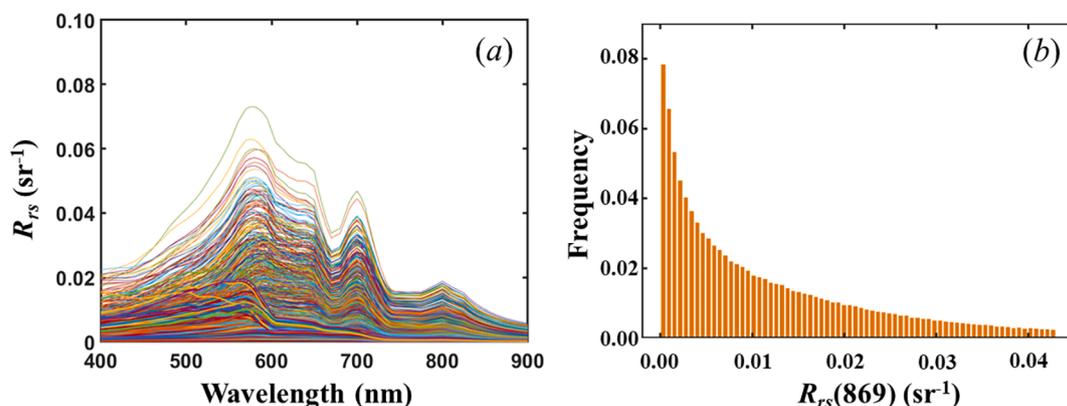


Fig. 3. (a) Examples of synthesized  $R_{rs}$  spectra, (b) Statistical distributions of all synthetic  $R_{rs}(869)$ .

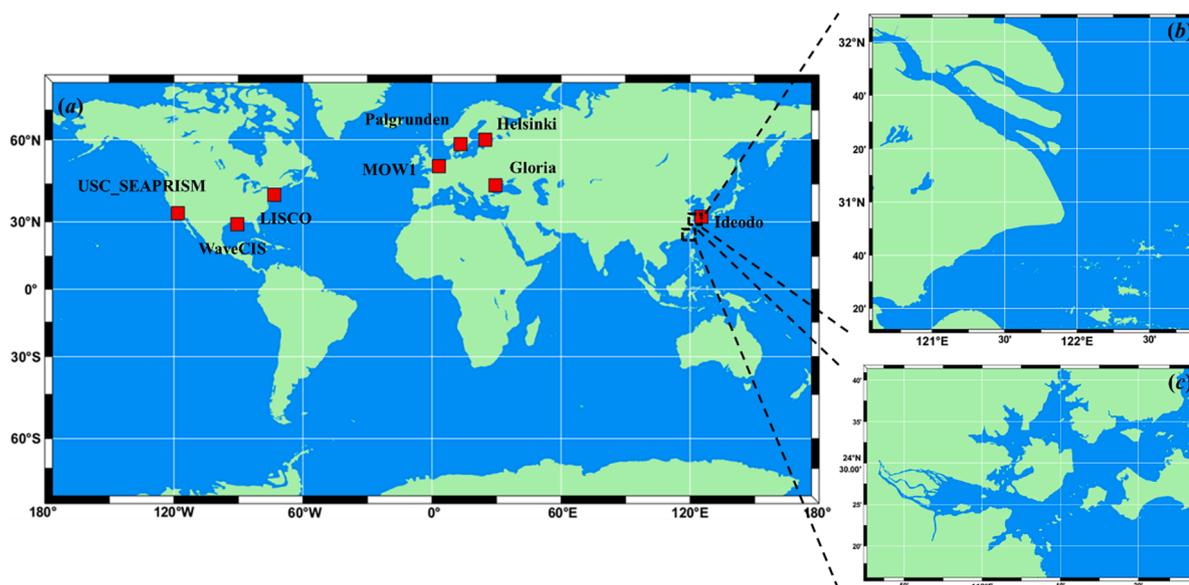


Fig. 4. (a) Locations of the AERONET-OC sites (red squares) used in this study. The black squares ((b) for the East China Sea and (c) for the Xiamen Bay area) show the coastal areas where the field measurements were obtained in the period of 2004–2018. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the literature (IOCCG-OC4G 2003; IOCCG 2006; Lee et al. 1999; Lee et al. 2016), some of the components and synthesizing steps are included in Supplementary Information available online for quick reference. The following lists key information associated with this dataset used for NN-EBVR training and evaluation.

- (1) The synthetic dataset consists of 300,000 sets of  $R_{rs}$  spectra, for wavelengths in the 400–900 nm domain with 1 nm interval; the ranges of  $a(440)$  and  $b_b(440)$  are  $0.0074\text{--}15.0\text{ m}^{-1}$  and  $0.0031\text{--}5.0\text{ m}^{-1}$ , respectively; and the resulted  $R_{rs}(869)$  is in a range of  $2.6 \times 10^{-6}\text{--}0.043\text{ sr}^{-1}$ , with  $R_{rs}(550)$  in a range of  $\sim 3.9 \times 10^{-5}\text{--}0.066\text{ sr}^{-1}$ . Fig. 3a shows examples of the simulated  $R_{rs}$  spectra, with Fig. 3b showing the histogram of  $R_{rs}(869)$ , indicating a sizable portion ( $\sim 25.9\%$ ) of the data is highly turbid waters ( $R_{rs}(869) > 0.02\text{ sr}^{-1}$ ).
- (2) For the evaluation of NN-EBVR, in addition to the 20% of the synthetic dataset, we also used 243 sets of  $R_{rs}$  spectra measured in the Yangtze River Estuary (YRE) and JiuLong River (JLR) (see black rectangles in Fig. 4) from twenty-two field campaigns during 2004–2018 to validate the performance of NN-EBVR. For these  $R_{rs}$  spectra, each  $R_{rs}(\lambda, 350\text{--}1050\text{ nm})$  was obtained by a hand-held spectroradiometer (GER-1500) following the standard

protocol (Mueller et al. 2000) conducted from the field campaigns. For the correction of surface reflectance, we adopted a method proposed by Lee et al. (2010), where a second-order correction was carried out through spectral optimization.

The 300,000 synthetic  $R_{rs}$  spectra are divided randomly by the 8:2 ratio, with 240,000 for training NN-EBVR, and 60,000 for the validation of NN-EBVR. For the training of NN-EBVR, to mimic the error-bearing  $R_{rs}$  of turbid waters in the initial round of  $ACA_{NIR-NN}$  (where  $R_{rs}(869)$  is assumed as 0 for all waters), the following was calculated for the synthetic and ground-measured  $R_{rs}$ ,

$$R'_{rs}(\text{visible}) = R_{rs}(\text{visible}) - R_{rs}(869) \tag{4}$$

and it is  $R'_{rs}$  used as the inputs to estimate both  $R_{rs}(748)$  and  $R_{rs}(869)$  in NN-EBVR.

### 3.2. AERONET-OC and MODISA matchup for validating $ACA_{NIR-NN}$

In addition to the evaluation of NN-EBVR, the ground-measured data from the eight AERONET-OC sites (USC\_SEAPRISM, Palgrunden, Gloria, Helsinki\_Lighthouse (Helsinki), WaveCIS\_Site\_CSI (WaveCIS), LISCO, Zeebrugge-MOW1 (MOW1) and Ideodo; see Fig. 4 for locations) during

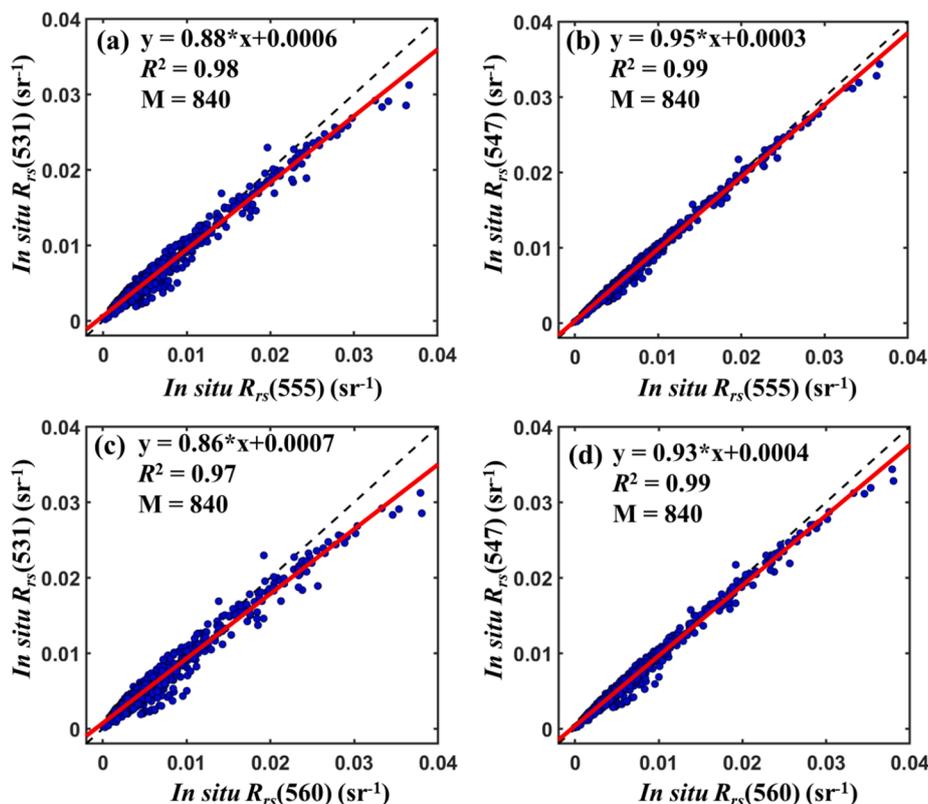


Fig. 5. Scatterplots between  $R_{rs}(555)$  and  $R_{rs}(531)$  (or  $R_{rs}(547)$ ) (upper panel), between  $R_{rs}(560)$  and  $R_{rs}(531)$  (or  $R_{rs}(547)$ ) (lower panel), from *in situ* hyperspectral  $R_{rs}$  data in Lee et al. (2016). M refers to all hyperspectral  $R_{rs}$  dataset in Lee et al. (2016).

2016–2017 were employed to validate the performance of  $ACA_{NIR-NN}$ . The quality-controlled Level 2.0 normalized  $L_w$  were downloaded from the AERONET-OC website (<http://aeronet.gsfc.nasa.gov/>), which are converted to  $R_{rs}$  through the following formula:

$$R_{rs}(\lambda) = \frac{L_{wn}(\lambda)}{F_0(\lambda)} \quad (5)$$

To determine the matchups between AERONET-OC and MODISA, we adopted the following criteria following Bailey and Werdell (2006): (1) the maximum time difference between the MODISA and AERONET-OC measurements is within  $\pm 1.5$  h. If there are two or more AERONET-OC measurements that meet this time difference, we linearly interpolated these AERONET-OC data to the overpass time of the satellite; (2) If the following two criteria are satisfied, i.e., there are more than 50% of valid pixels for satellite-retrieved  $R_{rs}$  within a  $3 \times 3$  box centered on the location of each AERONET-OC site and the coefficient of variation (CV) at each wavelength is less than 0.15, the averaged value of these valid pixels were used to compare with *in situ*  $R_{rs}$ . After applying the above criteria, a total of 138 matchups between the eight AERONET-OC sites and MODISA images during 2016–2017 were assembled.

To ensure the consistency of evaluation, the AERONET-OC  $R_{rs}$  bands must be adjusted to the MODISA bands in order to eliminate the impact due to the difference in center wavelengths between AERONET-OC (555 or 560 nm) and MODISA (531 and 547 nm). Similarly to the method in Wang et al. (2021a), we used linear relationships for four-band pairs of  $R_{rs}$  (555 nm vs 531 or 547 nm; 560 nm vs 531 or 547 nm; see Fig. 5) based on the same *in situ* dataset to establish conversion relationships between these  $R_{rs}$ . Subsequently, the measurements from AERONET-OC were converted to equivalent MODISA  $R_{rs}$  for the evaluation of  $ACA_{NIR-NN}$ .

### 3.3. Satellite data

In this study, we apply  $ACA_{NIR-NN}$  to process MODIS Aqua images over coastal areas to demonstrate its effectiveness. MODIS Aqua Level-1A images corresponding to the ground measurements were obtained from NASA’s OceanColor Web (<https://oceancolor.gsfc.nasa.gov/>), which is supported by the Ocean Biology Processing Group (OBPG) at NASA’s Goddard Space Flight Center. In addition to using  $ACA_{NIR-NN}$ , these MODISA images were also processed to Level-2 products using the  $ACA_{NIR-bio}$  and  $ACA_{NIR-SWIR}$  that are embedded in the SeaDAS (version 7.5.3) software package. After these processes, we eliminated the low-quality  $R_{rs}(\lambda)$  retrievals from all three ACAs where the  $l2\_flag$  HILT (very high or saturated observed radiance), LAND (land pixel), ATM-FAIL (atmospheric correction failure), HIGLINT (strong sun glint contamination) and CLDICE (probable cloud or ice contamination) occurred. The quality-controlled  $R_{rs}$  at bands 412, 443, 488, 531, 547 and 667 nm from the three ACAs were then compared with each other and ground measurements to assess the performances of these AC algorithms. Further, to demonstrate the performance of  $ACA_{NIR-NN}$  over coastal waters for sensors without SWIR bands, we selected one SeaWiFS image as an example, with data also downloaded from NASA’s OceanColor Web. The  $R_{rs}$  retrievals at 412, 443, 490, 555 and 670 nm from  $ACA_{NIR-NN}$  were compared with that from the  $ACA_{NIR-bio}$ . Note that since there are no SWIR bands for SeaWiFS, we used the  $R_{rs}$  retrievals from the same-day MODISA data (the difference in overpass time between the same-day SeaWiFS and MODISA collections is less than 0.5 h) retrieved by  $ACA_{NIR-SWIR}$  as the reference.

### 3.4. Statistical metrics for $ACA_{NIR-NN}$ and NN-EBVR

For the statistical evaluations, the performance of each AC approach was evaluated using five statistics parameters, including the coefficient of determination ( $R^2$ ) in linear regression analysis, Root Mean Square

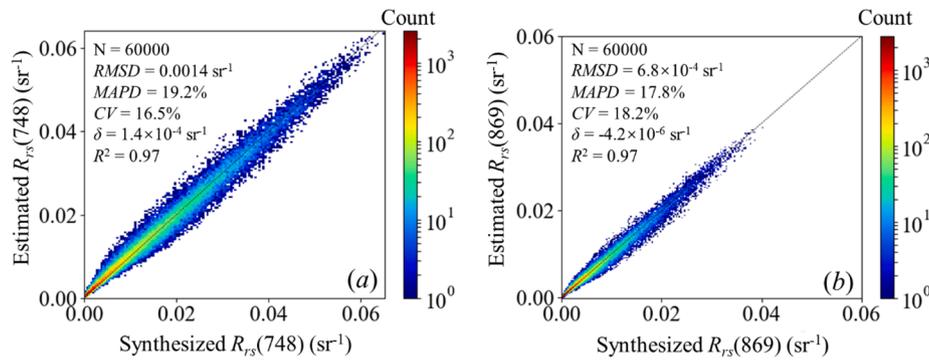


Fig. 6. Scatterplots between NN-EBVR estimated  $R_{rs}$ (NIR) and synthesized  $R_{rs}$ (NIR) for the 20% synthetic dataset. The color in the scatterplot represents the count of data points.

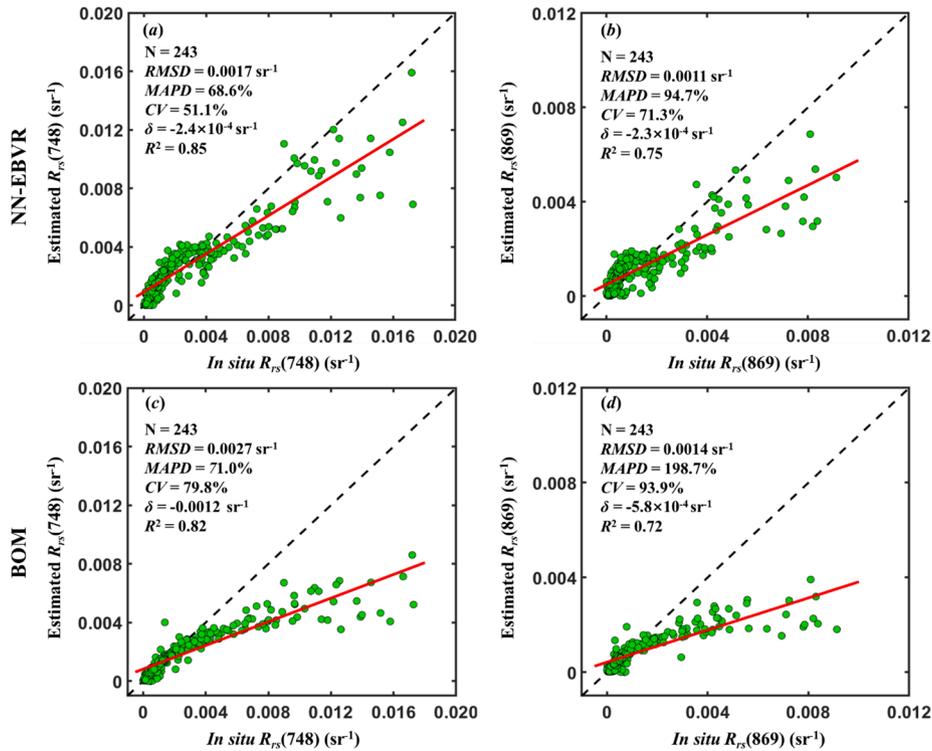


Fig. 7. Scatterplots compare the estimated  $R_{rs}$ (NIR) from NN-EBVR ((a) and (b)) and the estimated  $R_{rs}$ (NIR) from the bio-optical models adopted in  $ACA_{NIR-bio}$  ((c) and (d)) with the true *in situ* measurements. The black dash line represents the 1:1 line, the red solid line represents the linear regression corresponding to all evaluated datasets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Difference (*RMSD*), Mean Absolute Percentage Difference (*MAPD*), Coefficient of Variation and bias ( $\delta$ ). They are defined as follows ( $N$  is the number of data points):

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (Q_{est,i} - Q_{mea,i})^2}{N}} \quad (6a)$$

$$MAPD = \frac{1}{N} \sum_{i=1}^N \frac{|Q_{mea,i} - Q_{est,i}|}{Q_{mea,i}} \times 100\% \quad (6b)$$

$$CV = \frac{RMSD}{\text{Mean}(Q_{mea})} \times 100\% \quad (6c)$$

$$\delta = \frac{1}{N} \sum_{i=1}^N (Q_{est,i} - Q_{mea,i}) \quad (6d)$$

where  $Q_{est,i}$  and  $Q_{mea,i}$  are derived and known (*in situ*) values of  $R_{rs}$ , respectively.

Meanwhile, we also adopted similar metrics to validate the accuracy of estimated  $R_{rs}$ (NIR) from NN-EBVR, where the  $Q_{est,i}$  and  $Q_{mea,i}$  are estimated and synthesized (or *in situ*) values of  $R_{rs}$  at the two NIR bands, respectively.

## 4. Results and discussion

### 4.1. Validation of NN-EBVR

The accuracy of estimated  $R_{rs}$ (NIR) from error-bearing  $R_{rs}$ (visible) by NN-EBVR is first evaluated using the 20% synthetic dataset, with results for MODISA spectral settings shown in Fig. 6a–b as examples. It should be noted that although the inputs to NN-EBVR is the error-bearing  $R_{rs}$ (visible) ( $R'_{rs}$ ), it is the true  $R_{rs}$ (NIR) used to compare with the estimated  $R_{rs}$ (NIR) from NN-EBVR. Generally, for these validation

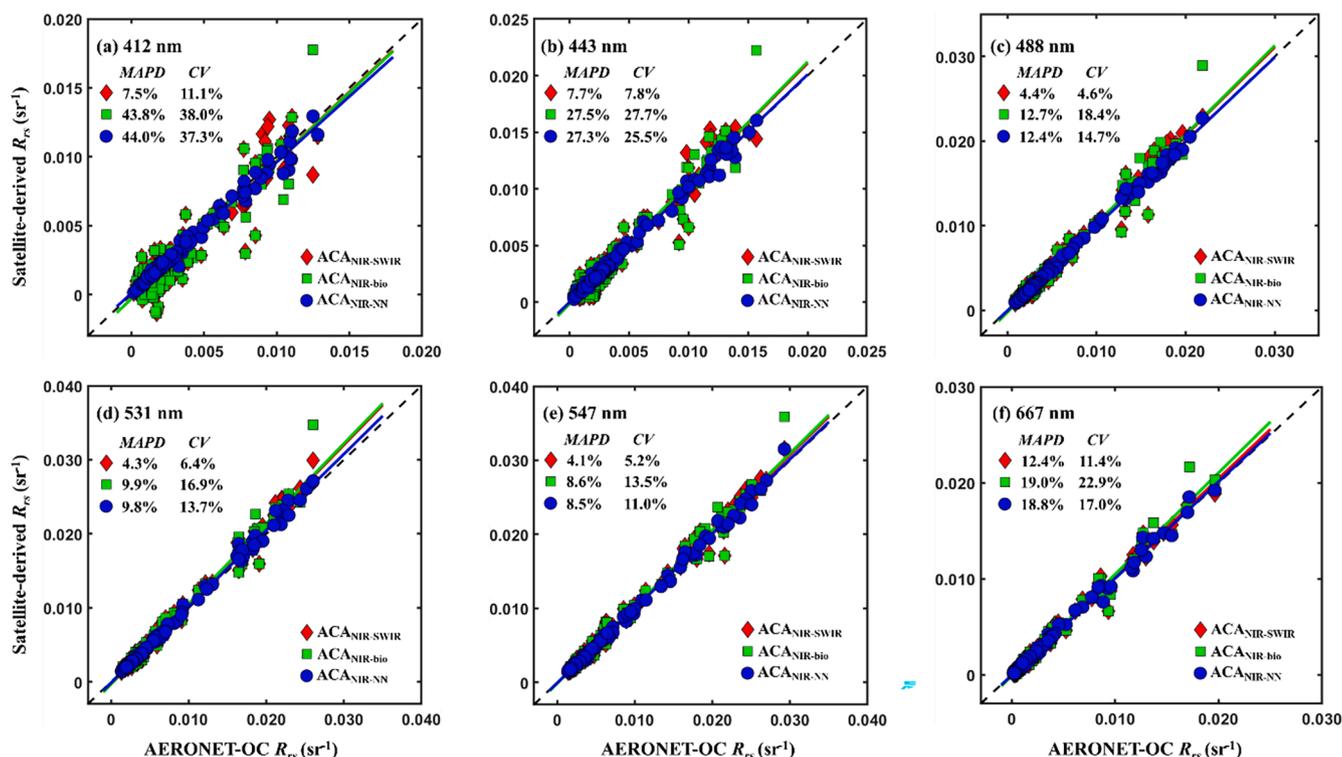


Fig. 8. Scatterplot comparison between MODISA-retrieved  $R_{rs}$  and *in situ*  $R_{rs}$  at bands 412, 443, 488, 531, 547 and 667 nm. The black dash line represents the 1:1 line, the colored solid lines represent the linear regression of the different AC algorithms (red line for  $ACA_{NIR-SWIR}$ , green line for  $ACA_{NIR-bio}$  and blue line for  $ACA_{NIR-NN}$ ). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

datasets, the values of  $R^2$  for these two NIR bands (748 nm and 869 nm) are both close to  $\sim 0.97$ , with values of *MAPD* both less than 20% and bias close to 0 (Fig. 6). We also compared the performance of NN-EBVR using data from ship-based *in situ* measurements (243 sets of data from twenty-two field campaigns during 2004–2018), with scatterplots and statistics shown in Fig. 7a–7b. Compared to using the 20% of the synthetic dataset, *MAPD* of estimated  $R_{rs}(748)$  and  $R_{rs}(869)$  from NN-EBVR with *in situ* measurements changed to 68.6% and 94.7%, respectively. It appears that most of the differences happened to waters with *in situ*  $R_{rs}(748)$  and  $R_{rs}(869)$  larger than  $0.004 \text{ sr}^{-1}$  and  $0.002 \text{ sr}^{-1}$ , respectively, suggesting that there are still rooms for NN-EBVR to improve the

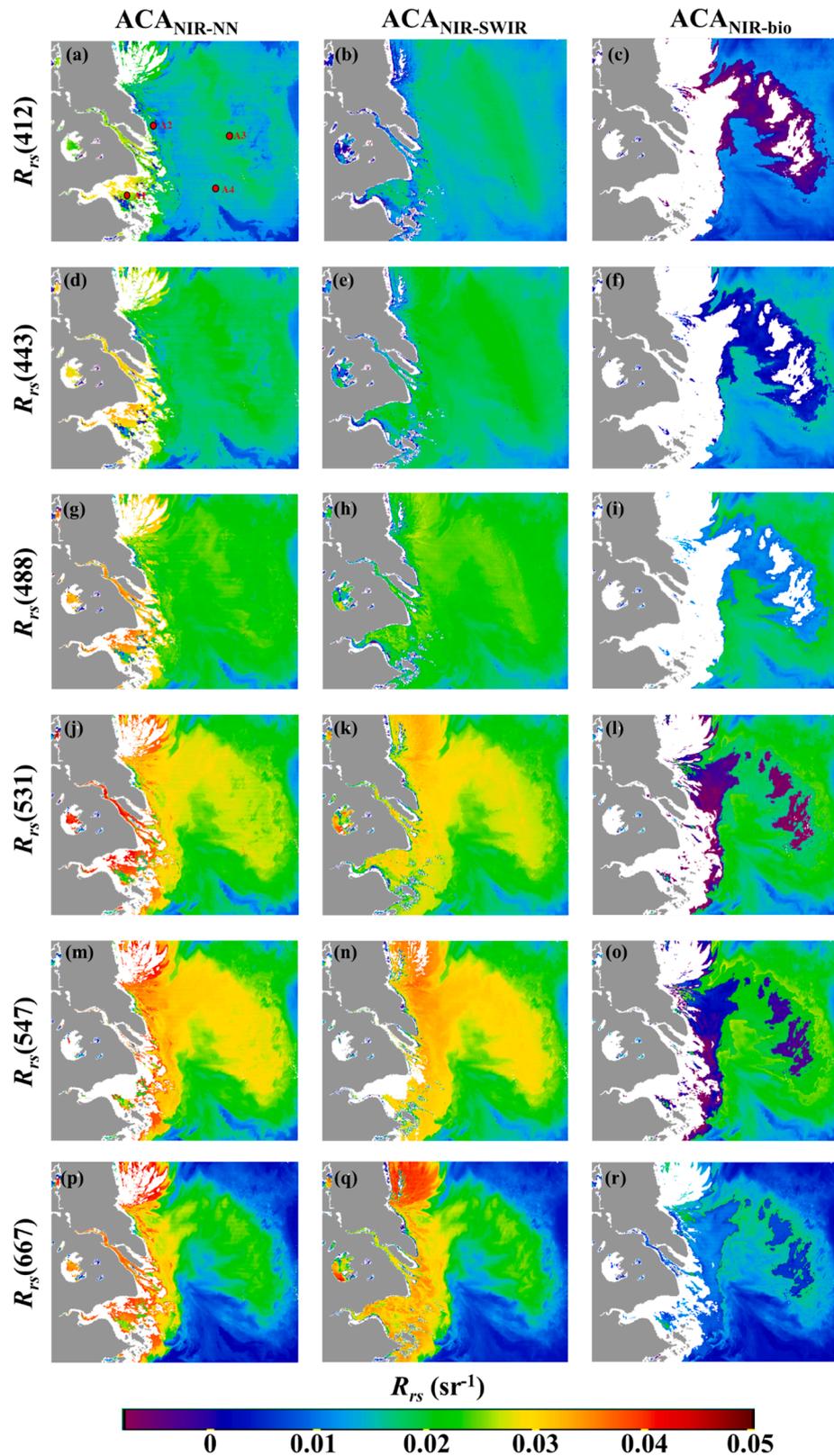
estimation of  $R_{rs}(748)$  and  $R_{rs}(869)$  in highly turbid waters. On the other hand, this larger difference could be the result of incomplete correction of the residual sun or sky glint in the field measurements of  $R_{rs}$ , which is difficult to accurately remove for coastal turbid waters with the above-water approach (Kutser et al. 2013; Lee et al. 2010).

On the other hand, assuming a water body with true  $R_{rs}(869)$  as  $0.008 \text{ sr}^{-1}$  (about the maximum value within the 243 sets of ship-based *in situ* measurements), for a common aerosol optical depth at 869 nm of 0.14, the solar zenith angle of  $30^\circ$ , sensor view zenith angle as  $30^\circ$ , and the standard atmospheric properties, the relative contribution of  $R_{rs}(869)$  to  $\rho_t(869)$  is about 52%. For such situations, it is found that

Table 1

Statistics of the accuracy measures for  $ACA_{NIR-NN}$  and two other atmospheric correction algorithms for six MODISA bands, as gauged using *in situ*  $R_{rs}$  measurements.

	Band	N	Slope	$R^2$	$RMSD (\times 10^{-4}) (\text{sr}^{-1})$	<i>MAPD</i> (%)	Bias $(\times 10^{-4}) (\text{sr}^{-1})$	<i>CV</i> (%)
$ACA_{NIR-NN}$	412	138	0.95	0.98	3.8	7.5	-0.70	11.1
	443	138	1.00	0.99	3.2	7.7	0.17	7.8
	488	138	0.99	0.99	2.7	4.4	-0.39	4.6
	531	138	1.03	0.99	4.5	4.3	-0.86	6.4
	547	138	1.01	0.99	3.8	4.1	-0.06	5.2
	667	138	1.01	0.99	3.2	12.4	-0.08	11.4
$ACA_{NIR-bio}$	412	138	1.00	0.86	13	43.8	-3.15	38.0
	443	138	1.07	0.94	11	27.5	1.19	27.7
	488	138	1.05	0.97	11	12.7	0.76	18.4
	531	138	1.08	0.98	12	9.	2.78	16.9
	547	138	1.04	0.98	9.8	8.6	1.36	13.5
	667	138	1.06	0.98	6.4	19.0	0.69	22.9
$ACA_{NIR-SWIR}$	412	138	1.00	0.86	13	44.0	-2.92	37.3
	443	138	1.06	0.94	10	27.3	1.23	25.5
	488	138	1.04	0.98	8.6	12.4	0.68	14.7
	531	138	1.08	0.99	9.5	9.80	2.63	13.7
	547	138	1.03	0.99	8.0	8.5	1.21	11.0
	667	138	1.02	0.99	4.8	18.8	0.38	17.0



**Fig. 9.** Spatial distributions of  $R_{rs}$  (412, 443, 488, 531, 547 and 667 nm) from the  $ACA_{NIR-NN}$  (left),  $ACA_{NIR-SWIR}$  (middle) and  $ACA_{NIR-bio}$  (right) over the East China Sea for MODISA image on April 7th, 2013. Grey color for land, white color for pixels flagged out by SeaDAS. The red dots are the locations of a few randomly selected pixels (termed as stations A1-A4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

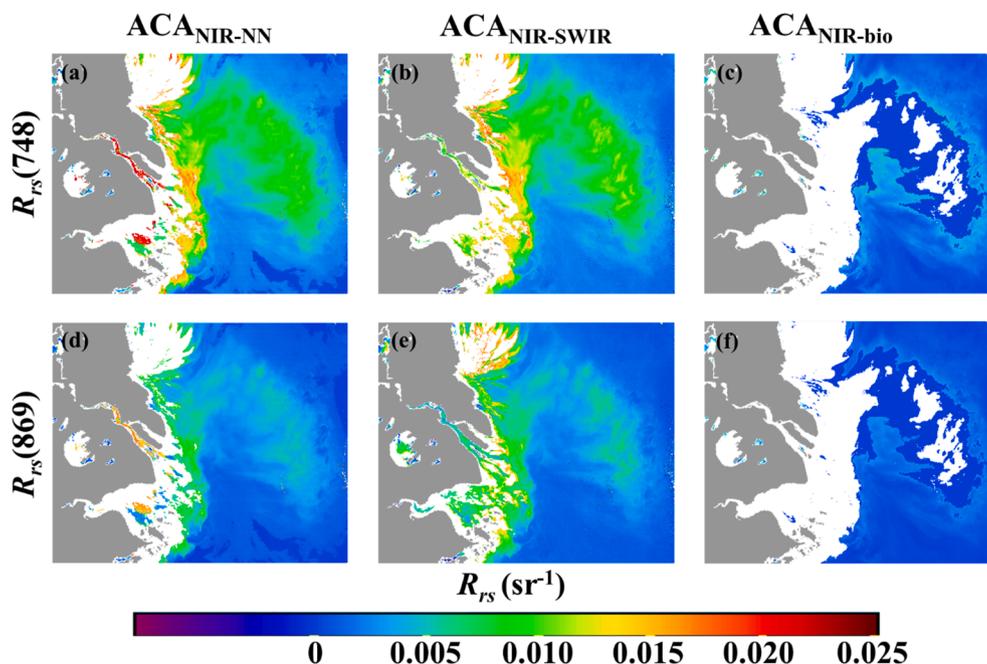


Fig. 10. Same as Fig. 9, but for  $R_{rs}$  at 748 and 869 nm.

uncertainty of 50% at  $R_{rs}(869)$  causes uncertainty for all  $R_{rs}(\text{visible})$  being less than 10%, which, as evidenced later, may not significantly impact the AC processes.

We also compared the estimated  $R_{rs}(\text{NIR})$  using the bio-optical models adopted in the conventional  $\text{ACA}_{\text{NIR-bio}}$  scheme with results shown in Fig. 7c–7d. Note that the inputs for the bio-optical model are also the error-bearing  $R_{rs}(\text{visible})$  as that in the NN-EBVR scheme. It is found that the estimated  $R_{rs}(\text{NIR})$  from this bio-optical model are in general significantly lower than *in situ* measurements, where for this dataset the *MAPDs* and *CVs* for both wavelengths are 71.0% and 79.8% at 748 nm, and 198.7% and 93.9% at 869 nm, respectively (see Fig. 7c–7d). However, for this dataset, it appears that BOM showed good performance for  $R_{rs}(748) < \sim 0.003 \text{ sr}^{-1}$  or  $R_{rs}(869) < \sim 0.0015 \text{ sr}^{-1}$ .

#### 4.2. Performance of $\text{ACA}_{\text{NIR-NN}}$ in processing satellite ocean color measurements

##### 4.2.1. Evaluation using AERONET-OC measurements

For the MODISA measurements described in Section 3, the performance of  $\text{ACA}_{\text{NIR-NN}}$  was assessed using data from eight AERONET-OC sites (a total of 138 matchups from 2016 to 2017) covering various coastal waters dominated by Chl or sediments (Zibordi et al. 2009) or dominated strongly by CDOM (AERONET-OC site “Helsinki”) (Zibordi et al. 2021). Fig. 8 shows MODISA-retrieved  $R_{rs}$  compared with *in situ*  $R_{rs}$  at bands 412, 443, 488, 531, 547 and 667 nm. For the 138 matchups, the satellite-retrieved  $R_{rs}$  from  $\text{ACA}_{\text{NIR-NN}}$  (termed as  $R_{rs,\text{NIR-NN}}$ ) at each visible band matched very well with *in situ* measurements ( $R_{rs,\text{in situ}}$ ), with  $R^2$  about  $\sim 0.98$  and *RMSD* varying from  $2.7 \times 10^{-4} \text{ sr}^{-1}$  to  $4.5 \times 10^{-4} \text{ sr}^{-1}$  and *MAPD* varying from 4.1% to 12.4% (see Table 1). Specifically, due to the maximum  $R_{rs}$  for most matchup *in situ* measurements being generally around 550 nm, the retrieved  $R_{rs,\text{NIR-NN}}$  has the best accuracy at 488, 531 and 547 nm, with *MAPD* of 4.4%, 4.3% and 4.1%, respectively. For accuracy of  $R_{rs,\text{NIR-NN}}$  retrievals at 412 and 443 nm, compared with *in situ*  $R_{rs}$  measurements, the  $\text{ACA}_{\text{NIR-NN}}$  slightly overestimated (7.7%)  $R_{rs}(443)$  (see Fig. 8b), while slightly underestimated (7.5%, see Fig. 8a)  $R_{rs}(412)$ , where such kind of differences suggest highly consistent determination of  $R_{rs}$  between MODISA and AERONET-OC measurements.

We also assessed the performance of  $\text{ACA}_{\text{NIR-bio}}$  (its results are

termed as  $R_{rs,\text{NIR-bio}}$ ) and  $\text{ACA}_{\text{NIR-SWIR}}$  (its results are termed as  $R_{rs,\text{NIR-SWIR}}$ ) for the same matchup dataset, with results also included in Fig. 8 and statistical measures presented in Table 1. For  $\text{ACA}_{\text{NIR-bio}}$ , the lowest  $R^2$  value is 0.86 while the highest *MAPD* is 43.8% (both occurred at 412 nm). Meanwhile,  $\text{ACA}_{\text{NIR-SWIR}}$  also exhibits the worst performance at 412 nm with  $R^2$  as 0.86 and *MAPD* as 44.0%. Note that  $\text{ACA}_{\text{NIR-SWIR}}$  is a combination of  $\text{ACA}_{\text{NIR-bio}}$  and  $\text{ACA}_{\text{SWIR}}$ , where  $\text{ACA}_{\text{SWIR}}$  is used for data processing only when the Turbid Water Index ( $T_{\text{ind}}$ ) meets the threshold (usually 1.3) (Wang and Shi 2007), otherwise, it is  $\text{ACA}_{\text{NIR-bio}}$  employed. For this dataset, with  $\text{ACA}_{\text{NIR-SWIR}}$ , only 19 matchups used  $\text{ACA}_{\text{SWIR}}$  while the other 119 matchups still used  $\text{ACA}_{\text{NIR-bio}}$ , therefore the statistics measurements for  $\text{ACA}_{\text{NIR-bio}}$  and  $\text{ACA}_{\text{NIR-SWIR}}$  are similar.

For such coastal waters, not surprisingly, the retrieved  $R_{rs,\text{NIR-bio}}$  and  $R_{rs,\text{NIR-SWIR}}$  show better agreement with the *in situ* measurements for wavelengths longer than 443 nm, with *MAPDs* being less than 28%. In general, for these AERONET-OC sites, it appears that  $\text{ACA}_{\text{NIR-bio}}$  and  $\text{ACA}_{\text{NIR-SWIR}}$  exhibit lower accuracy than  $\text{ACA}_{\text{NIR-NN}}$ , where the *MAPD* values of  $R_{rs}$  retrievals by  $\text{ACA}_{\text{NIR-bio}}$  and  $\text{ACA}_{\text{NIR-SWIR}}$  are larger by a factor of two or three than that by  $\text{ACA}_{\text{NIR-NN}}$  for each band, e.g., 44.0% vs 7.5% at 412 nm, or 8.6% vs 4.1% at 547 nm (Table 1). For the band at 412 nm, although these three AC algorithms generally underestimate  $R_{rs}(412)$  retrievals, the bias of  $R_{rs,\text{NIR-bio}}(412)$  and  $R_{rs,\text{NIR-SWIR}}(412)$  is larger than that of  $R_{rs,\text{NIR-NN}}(412)$ , which are  $-3.15 \times 10^{-4} \text{ sr}^{-1}$  and  $-2.92 \times 10^{-4} \text{ sr}^{-1}$ , respectively (Table 1), even sometimes negative retrievals occurring for  $R_{rs,\text{NIR-bio}}(412)$  and  $R_{rs,\text{NIR-SWIR}}(412)$ , a result not shown from  $\text{ACA}_{\text{NIR-NN}}$ . On the other hand, it is necessary to keep in mind that these AERONET-OC sites do not cover extremely turbid waters, how  $\text{ACA}_{\text{NIR-NN}}$  behaves in such environments remains to be seen.

##### 4.2.2. Evaluation using MODISA data over turbid waters

Since the locations of the above AERONET-OC sites do not cover highly turbid waters, we selected Yangtze River Estuary and adjacent coastal waters in the East China Sea, a well-known environment of extremely turbid waters (Wang et al. 2007), to further evaluate the performance of  $\text{ACA}_{\text{NIR-NN}}$ . In this process, we used  $R_{rs}$  results from  $\text{ACA}_{\text{NIR-SWIR}}$  as the reference due to no concurrent field measurements. For the MODISA image obtained on April 7, 2013, Figs. 9 and 10 shows the spatial distribution of  $R_{rs}$  at 412, 443, 488, 531, 547, 667, 745 and 869 nm generated by  $\text{ACA}_{\text{NIR-NN}}$  (left column), which also included  $R_{rs}$

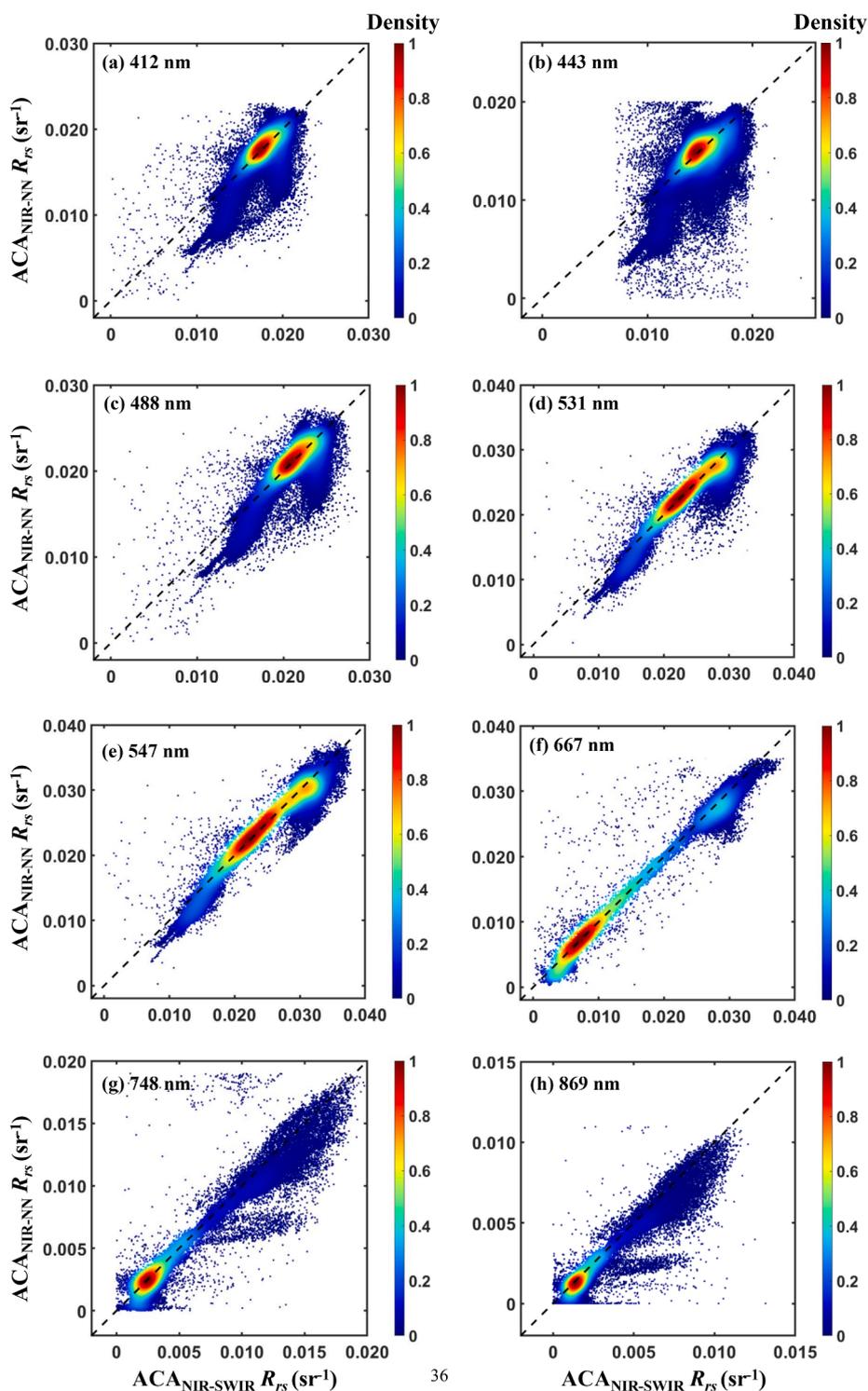


Fig. 11. Scatterplot comparisons of  $R_{rs}$  at 412, 443, 488, 531, 547, 667, 748 and 869 nm between retrievals from  $ACA_{NIR-SWIR}$  (x-axis) and  $ACA_{NIR-NN}$  (y-axis), for the MODISA image on April 7th, 2013 over the East China Sea. The color in the scatterplots represents the density of the points.

generated by  $ACA_{NIR-SWIR}$  (middle column) and  $ACA_{NIR-bio}$  (right column) for comparison. Fig. 11 shows scatterplots of  $R_{rs}$  between  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$ , with detailed statistics presented in Table 2. Overall, although there are some differences in  $R_{rs}$  values at each band between  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$ , the retrieved  $R_{rs}$  from the two schemes are highly consistent in the visible ( $MAPD$  values generally <9.8%) (see Table 2). Again, due to the maximum  $R_{rs}$  generally appearing in the green–red domain for these turbid waters, the best agreement between  $R_{rs,NIR-NN}$  and  $R_{rs,NIR-SWIR}$  is found at the green and

red bands (Fig. 11d–11f), with  $MAPDs$  of 4.2%, 4.0% and 6.6% and  $RMSDs$  of  $0.0014\text{ sr}^{-1}$ ,  $0.0014\text{ sr}^{-1}$  and  $0.0010\text{ sr}^{-1}$ , respectively.

In addition to the overall consistency of satellite-retrieved  $R_{rs}$  between  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$ , the spatial coverage of valid data is also an important metric. It is found that,  $R_{rs}$  retrievals at each band derived from  $ACA_{NIR-NN}$  are absent in a portion of Subei Shoal (SBS), the mouth of Yangtze River Estuar, Hangzhou Bay (HZB) and Taihu Lake. These absence of  $R_{rs}$  from  $ACA_{NIR-NN}$  are results of saturated signals (resulted in invalid data) at the two MODISA NIR bands caused by highly

**Table 2**

Same as Table 1, but for comparison between  $R_{rs}$  retrieved by  $ACA_{NIR-SWIR}$  (x-axis) and  $ACA_{NIR-NN}$  or  $ACA_{NIR-bio}$  (y-axis) for MODISA image on April 7th, 2013 over the East China Sea.

	Band	Slope	$R^2$	$RMSD$ ( $sr^{-1}$ )	$MAPD$ (%)	$CV$ (%)
$ACA_{NIR-NN}$	412	0.93	0.55	0.0022	9.8	13.6
	443	0.99	0.58	0.0020	7.7	11.0
	488	0.97	0.72	0.0017	5.4	8.1
	531	0.97	0.91	0.0014	4.2	5.8
	547	0.98	0.95	0.0014	4.0	5.1
	667	0.98	0.98	0.0010	6.6	5.4
	748	0.90	0.94	$9.6 \times 10^{-4}$	22.6	25.4
	869	0.87	0.91	$6.5 \times 10^{-4}$	19.3	26.6
$ACA_{NIR-bio}$	412	0.82	0.51	0.0045	31.6	33.2
	443	0.76	0.35	0.0043	23.6	24.5
	488	0.74	0.40	0.0034	14.1	14.3
	531	1.05	0.73	0.0022	13.1	12.8
	547	0.95	0.82	0.0021	12.6	12.2
	667	1.03	0.96	0.0013	13.7	14.6
	748	-0.12	0.95	0.0509	904.3	953.8
	869	-0.05	0.93	0.0481	70.7	77.3

turbid waters, consequently no  $R_{rs}$  at other visible bands could be retrieved by  $ACA_{NIR-NN}$ . In contrast, because  $ACA_{NIR-SWIR}$  used data in the SWIR bands to carry out atmosphere correction and there are valid data at the SWIR bands over these coastal areas, there are valid  $R_{rs,NIR-SWIR}$  retrievals in the visible bands for these areas (except 547 nm band), although no valid data either at the two NIR bands. For  $R_{rs}(547)$  (see Fig. 9m-9n), both  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$  have no valid data over most portion of HZB, due to that the values of  $R_{rs}$  from these two ACAs reached the maximum threshold ( $\sim 0.12 sr^{-1}$ ) set in SeaDAS (version 7.5.3).

The above comparisons are further highlighted using  $R_{rs}$  spectra at four sites (see red dots Fig. 9a) covering various turbidities, with A1 in HZB for extremely turbid water, while A4 in the East China Sea for moderate turbid water. The averaged  $R_{rs}$  spectra within a  $3 \times 3$  window centered at the four locations are shown in Fig. 12. Generally, except A1, the  $R_{rs}$  spectra retrieved by  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$  match each other very well. For location A1 (see Fig. 12a), as described above, there are no valid  $R_{rs}$  at 547 nm due to that it reached the maximum threshold of  $R_{rs}$

retrievals. On the other hand, the spectral shapes of  $R_{rs,NIR-NN}$  and  $R_{rs,NIR-SWIR}$  are quite similar, although the former is higher in values. Based on report in the literature (Chen et al. 2014; Pan et al. 2017), the value of  $R_{rs,NIR-NN}(869)$  for A1 is more reasonable.

Further, as presented in many studies (Jiang and Wang 2014; Rudnick et al. 2000; Shehhi et al. 2017; Xue et al. 2021), it is found that  $ACA_{NIR-bio}$  failed to produce  $R_{rs}$  for many turbid areas (see Figs. 9 and 10, right column). For some less turbid waters (the right portion of the image), although there are  $R_{rs}$  data products from  $ACA_{NIR-bio}$ , it is found that the values of  $R_{rs,NIR-bio}$  are significantly lower than that of  $R_{rs,NIR-SWIR}$  or  $R_{rs,NIR-NN}$ , even negative for some areas (see Fig. 9c, 9l and 9o). The above comparisons highlight a significant improvement of  $ACA_{NIR-NN}$  over  $ACA_{NIR-bio}$  for these turbid waters, although both used the same measurements as inputs for data processing.

In addition to comparing the magnitude and spatial distribution of  $R_{rs}$  retrievals from all three algorithms, we also counted the total number of iterations of  $ACA_{NIR-NN}$  and  $ACA_{NIR-bio}$  for processing an image with  $500 \times 400$  pixels. It is found that on average  $ACA_{NIR-NN}$  took 2–3 iterations for each pixel, while  $ACA_{NIR-bio}$  took 3–4 iterations, so the total number of iterations (391,302) of  $ACA_{NIR-NN}$  is a factor of two fewer than that (721,527) of  $ACA_{NIR-bio}$ , suggesting  $ACA_{NIR-NN}$  is also more efficient in data processing.

### 4.3. Application to SeaWiFS images

To further demonstrate the performance of  $ACA_{NIR-NN}$  to satellite sensors without SWIR bands, as an example, we used a SeaWiFS image over the mouth of the Amazon River and adjacent waters on June 25, 2003, which was processed by both  $ACA_{NIR-NN}$  (the NN-EBVR was revised to match wavelength band settings of SeaWiFS) and  $ACA_{NIR-bio}$ . Since the difference in overpass time between the same-day SeaWiFS and MODISA collections is less than 0.5 h, we used the  $R_{rs}$  products from the same-day MODISA image retrieved by  $ACA_{NIR-SWIR}$  as the reference (included in Fig. 13). The left (by  $ACA_{NIR-NN}$ ) and middle (by  $ACA_{NIR-bio}$ ) panels of Fig. 13 show the spatial distribution of  $R_{rs}$  at 412, 443, 490, 555 and 670 nm over the mouth of the Amazon River from the SeaWiFS measurements, while the right panel shows the spatial distribution of  $R_{rs}$  at similar wavelengths retrieved by  $ACA_{NIR-SWIR}$  of the same area from the concurrent MODISA measurements. Meanwhile, Fig. 14 shows

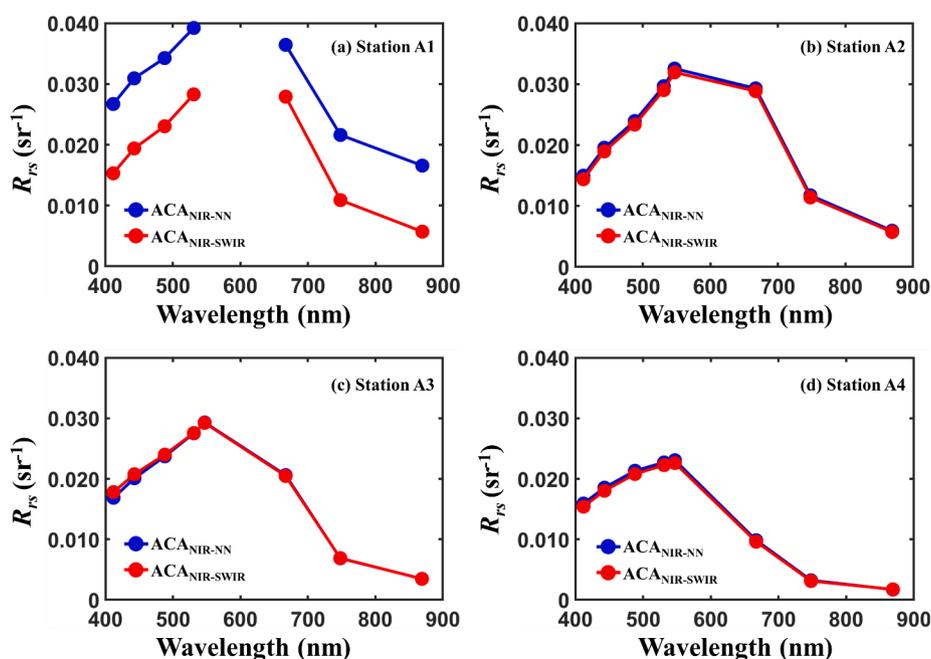
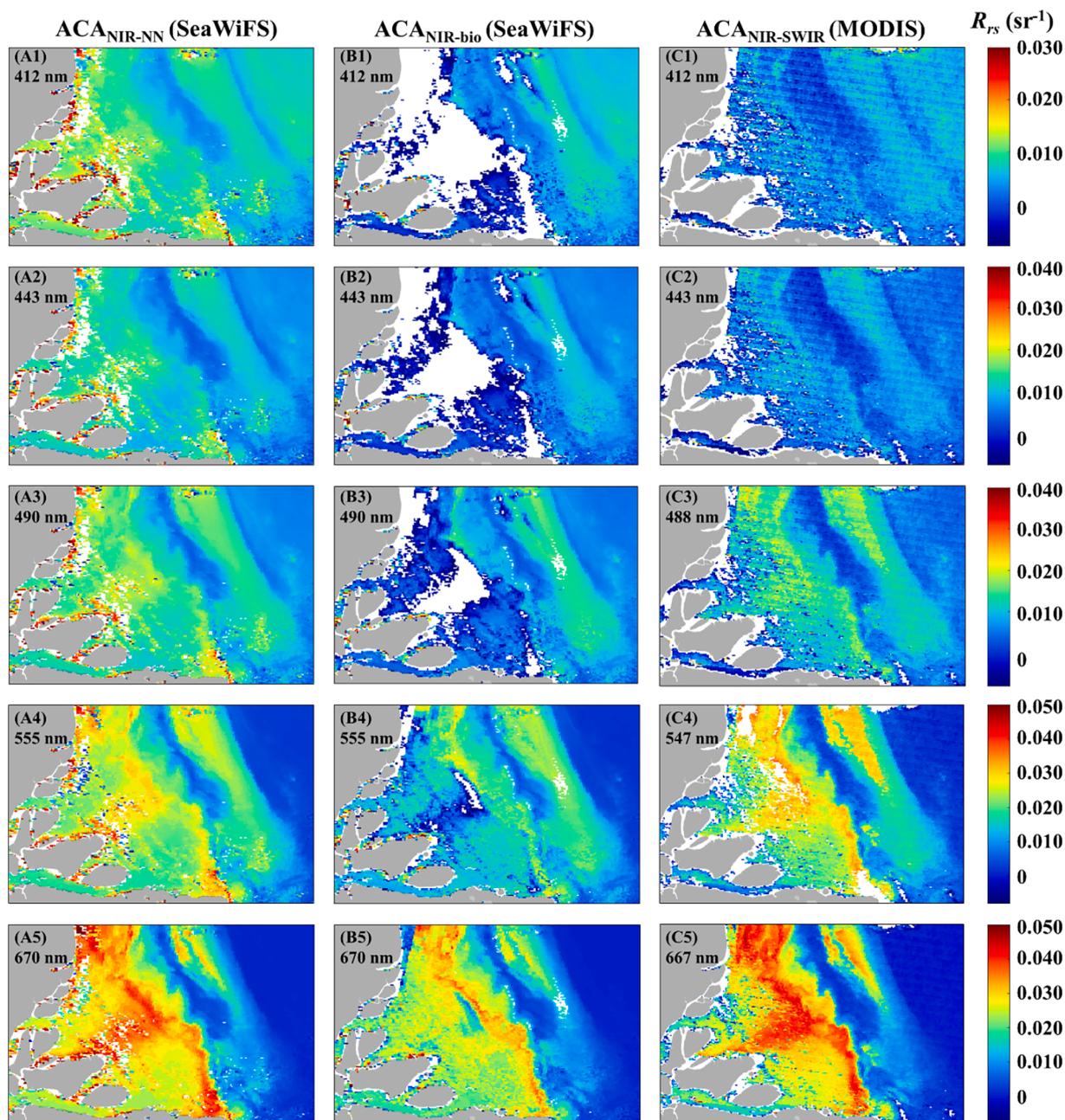


Fig. 12. Comparison of  $R_{rs}$  spectra from  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$  for the four selected locations in Fig. 9a.



**Fig. 13.** Spatial distributions of  $R_{rs}$  retrieved by  $ACA_{NIR-NN}$  (left panel),  $ACA_{NIR-bio}$  (middle panel), and  $ACA_{NIR-SWIR}$  (right panel) over the mouth of Amazon River and adjacent waters on June 25, 2003. The left and middle panels correspond to the SeaWiFS image, while the right panel corresponds to the concurrent MODISA image. Grey color for land, white color for pixels flagged out by SeaDAS (version 7.5.3).

scatterplots between  $R_{rs}$  from the SeaWiFS image by  $ACA_{NIR-NN}$  and  $R_{rs}$  from the MODISA image by  $ACA_{NIR-SWIR}$ .

It is found that for these turbid waters,  $R_{rs,NIR-NN}$  are very consistent with  $R_{rs,NIR-SWIR}$ , especially for the green–red bands, which have  $R^2$  values as high as  $\sim 0.93$ , a slope close to 1.0, and a near-zero bias, although they were obtained from two sensors with different AC schemes. Larger difference happened at the 412 nm band, where  $R_{rs,NIR-NN}$  from SeaWiFS are much higher than  $R_{rs,NIR-SWIR}$  from MODISA (see Fig. 13a), with the MAPD value as high as 246.3%. This larger difference is in part due to that, for many coastal pixels  $R_{rs,NIR-SWIR}$  from MODISA are negative (truncated in the scatterplots) or close to 0, but there are no negative  $R_{rs,NIR-NN}$  from SeaWiFS. It requires analysis of a large number of concurrent SeaWiFS and MODISA images in order to have a complete understanding of these differences, which is underway but out of the scope of this effort.

For the retrievals by the standard  $ACA_{NIR-bio}$ , the spatial patterns in general are similar to that of  $ACA_{NIR-NN}$  and  $ACA_{NIR-SWIR}$  for offshore waters (see Fig. 13, middle panel,  $R^2$  value as high as  $\sim 0.61$ ). However, over the mouth of Amazon River, the values of  $R_{rs,NIR-bio}$  are lower than both  $R_{rs,NIR-NN}$  and  $R_{rs,NIR-SWIR}$ . Especially for the blue bands, the retrieved  $R_{rs,NIR-bio}$  is negative for a larger portion of the image (see Fig. 13B1–13B3), which echoes findings in previous studies that  $ACA_{NIR-bio}$  has difficulties in processing highly turbid waters (Jiang and Wang 2014; Ruddick et al. 2000; Shehhi et al. 2017; Xue et al. 2021).

### 5. Conclusions

In ocean color remote sensing, it is still a challenge for atmospheric correction in turbid coastal areas where water-leaving radiance in the NIR (sometimes even SWIR) bands are no longer negligible due to high

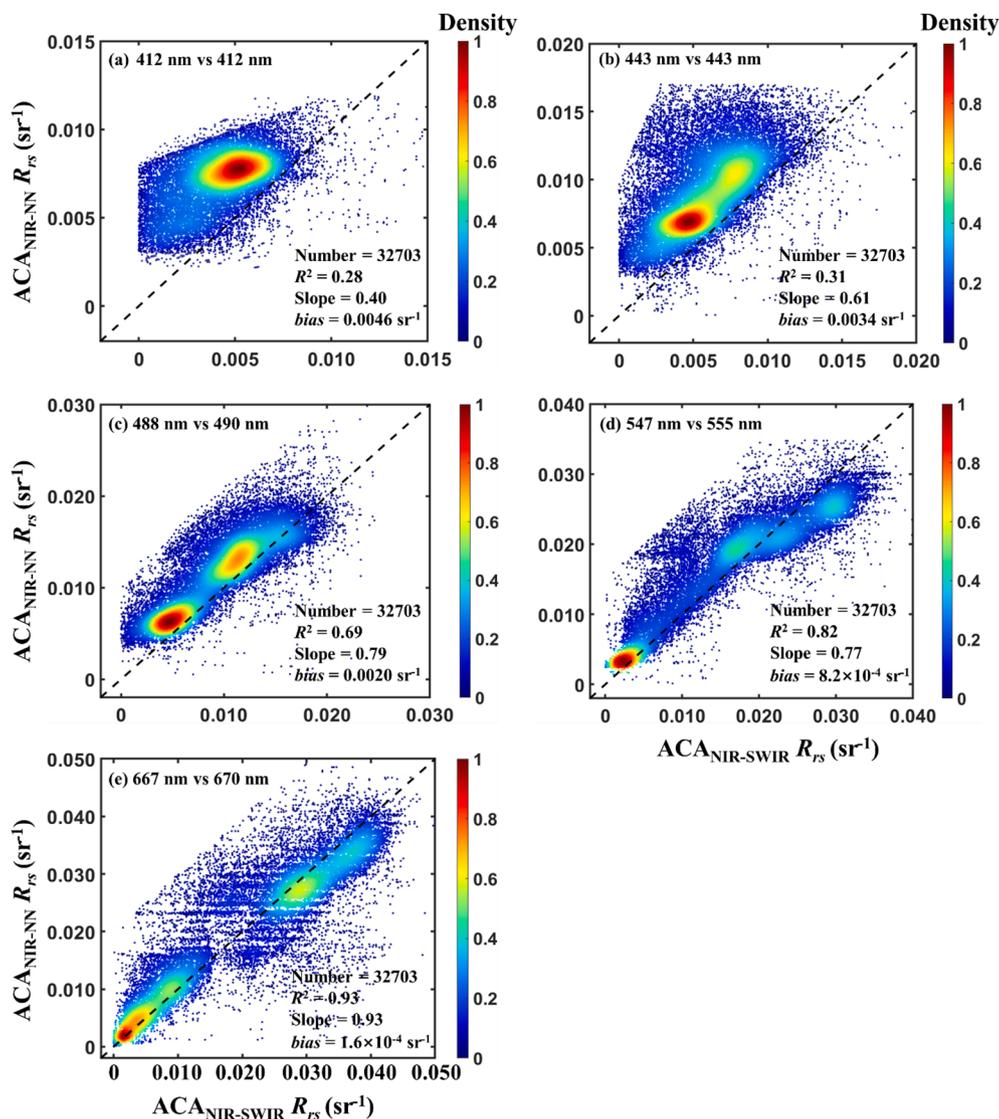


Fig. 14. Same as Fig. 11, but for results derived from  $ACA_{NIR-SWIR}$  (x-axis) and  $ACA_{NIR-NN}$  (y-axis) for June 25, 2003 MODIS (x-axis) and SeaWiFS (y-axis) images over the mouth of Amazon River and adjacent waters.

concentrations of suspended particulate matters. In this paper, based on SeaDAS (version 7.5.3), we present a scheme,  $ACA_{NIR-NN}$ , to estimate  $R_{rs}(NIR)$  for atmospheric correction, where  $R_{rs}(NIR)$  is estimated from available  $R_{rs}$  in the visible bands with specifically designed Neural Networks (NN-EBVR). The performance of  $ACA_{NIR-NN}$  was evaluated using MODISA measurements over eight different AERONET-OC sites. It is found that the *MAPD* of  $R_{rs}$  obtained from  $ACA_{NIR-NN}$  at each visible band is less than 12.4%, but the *MAPD* values of  $R_{rs}$  in the visible band obtained by  $ACA_{NIR-bio}$  could be a factor of two higher than that from  $ACA_{NIR-NN}$ . We further demonstrated the applicability of  $ACA_{NIR-NN}$  over turbid waters for MODISA (East China Sea) and SeaWiFS images (the mouth of Amazon River). The spatial patterns and values of  $R_{rs}$  retrievals by  $ACA_{NIR-NN}$  show similar characteristics to that of  $ACA_{NIR-SWIR}$ , even over highly turbid areas. These results demonstrated that  $ACA_{NIR-NN}$  (based on SeaDAS (version 7.5.3)) is a viable option to generate promising  $R_{rs}$  in turbid coastal waters for sensors having no SWIR bands. Further, because  $ACA_{NIR-NN}$  can also estimate  $R_{rs}(NIR)$  of clear waters, there is no need to have an artificial switch between clear and turbid waters in the data processing, therefore seamless  $R_{rs}$  data products can be generated when a satellite ocean color image covers both clear and turbid waters.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant number #41830102, #41941008, and #41890803) and the SanMing New Infrastructure Industry Development Limited Company (grant number #2021350204004224). The authors would like to thank NASA OBPG for providing the satellite data and SeaDAS software package. The colleagues managing the AERONET-OC stations worldwide are kindly thanked for their efforts in keeping the stations running and for the use of the data: Burton Jones and Matthew Ragan (USC-SEAPRISM), Susanne Kratzer (Palgrunden), Giuseppe Zibordi (Gloria), Giuseppe Zibordi (Helsinki Lighthouse), Sherwin Ladner (WaveCIS\_Site\_CSI), Sam Ahmed and Alex Gilerson (LISCO), Dimitry Van der Zande (Zeebrugge-MOW1) and Young-Je Park and Hak-Yeol You (Ieodo Station). We also thank Shaoling Shang and Gong Lin for sharing field measurements. Constructive comments and suggestions from four

anonymous reviewers are greatly appreciated.

## Appendix A. Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.isprsjrs.2022.10.014>.

## References

- Bailey, S.W., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Opt. Express* 18, 7521–7527.
- Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sens. Environ.* 102 (1–2), 12–23.
- Carswell, T., Costa, M., Young, E., Komick, N., Gower, J., Sweeting, R., 2017. Evaluation of MODIS-Aqua atmospheric correction and chlorophyll products of Western North American coastal waters based on 13 years of data. *Remote Sensing* 9, 1063.
- Chen, S., Zhang, T., Hu, L., 2014. Evaluation of the NIR-SWIR atmospheric correction algorithm for MODIS-Aqua over the Eastern China Seas. *Int. J. Remote Sens.* 35 (11–12), 4239–4251.
- Choi, K., Joo, D., Kim, J., 2017. Kapre: On-gpu audio preprocessing layers for a quick implementation of deep neural network models with keras. *arXiv preprint arXiv:1706.05781*.
- Chomko, R.M., Gordon, H.R., Maritorena, S., Siegel, D.A., 2003. Simultaneous retrieval of oceanic and atmospheric parameters for ocean color imagery by spectral optimization: a validation. *Remote Sens. Environ.* 84 (2), 208–220.
- Fan, Y., Li, W., Gatebe, C.K., Jamet, C., Zibordi, G., Schroeder, T., Stamnes, K., 2017. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sens. Environ.* 199, 218–240.
- Fan, Y., Li, W., Chen, N., Ahn, J.-H., Park, Y.-J., Kratzer, S., Schroeder, T., Ishizaka, J., Chang, R., Stamnes, K., 2021. OC-SMART: A machine learning based data analysis platform for satellite ocean color sensors. *Remote Sens. Environ.* 253, 112236.
- Gordon, H.R., Brown, J.W., Evans, R.H., 1988. Exact Rayleigh scattering calculations for use with Nimbus-7 Coastal Zone Color Scanner. *Appl. Opt.* 27, 862–871.
- Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over oceans with SeaWiFS: A preliminary algorithm. *Appl. Opt.* 33, 443–452.
- Goyens, C., Jamet, C., Schroeder, T., 2013. Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sens. Environ.* 131, 63–75.
- Hu, C., Lee, Z., Franz, B., 2012. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J. Geophys. Res.* 117, C01011.
- Ibrahim, A., Franz, B., Ahmad, Z., Healy, R., Knobelspiesse, K., Gao, B.-C., Proctor, C., Zhai, P.-W., 2018. Atmospheric correction for hyperspectral ocean color retrieval with application to the Hyperspectral Imager for the Coastal Ocean (HICO). *Remote Sens. Environ.* 204, 60–75.
- Ibrahim, A., Franz, B.A., Ahmad, Z., Bailey, S.W., 2019. Multiband atmospheric correction algorithm for ocean color retrievals. *Front. Earth Sci.* 7, 116.
- IOCCG, 2010. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. In: Wang, M. (Ed.), Reports of the International Ocean-Colour Coordinating Group. IOCCG, Dartmouth, Canada, p. 83.
- IOCCG, 2006. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. In Z.-P. Lee (Ed.), Reports of the International Ocean-Colour Coordinating Group, No. 5 (p. 126). Dartmouth, Canada: IOCCG.
- IOCCG-OCAG, 2003. Model, parameters, and approaches that used to generate wide range of absorption and backscattering spectra. In: International Ocean Colour Coordinating Group, [http://www.ioccg.org/groups/OCAG\\_data.html](http://www.ioccg.org/groups/OCAG_data.html).
- Jamet, C., Loisel, H., Kuchinke, C.P., Ruddick, K., Zibordi, G., Feng, H., 2011. Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements. *Remote Sens. Environ.* 115 (8), 1955–1965.
- Jiang, L., Wang, M., 2014. Improved near-infrared ocean reflectance correction algorithm for satellite ocean color data processing. *Opt. Express* 22, 21657–21678.
- Ketkar, N., 2017. Introduction to keras. In: Ketkar, N. (Ed.), *Deep Learning With Python*. Apress, Berkeley, CA, pp. 97–111.
- Kuchinke, C.P., Gordon, H.R., Harding, L.W., Voss, K.J., 2009. Spectral optimization for constituent retrieval in Case 2 waters II: Validation study in the Chesapeake Bay. *Remote Sens. Environ.* 113 (3), 610–621.
- Kutser, T., Vahtmäe, E., Paavel, B., Kauer, T., 2013. Removing glint effects from field radiometry data measured in optically complex coastal and inland waters. *Remote Sens. Environ.* 133, 85–89.
- Lavender, S.J., Pinkerton, M.H., Moore, G.F., Aiken, J., Blondeau-Patissier, D., 2005. Modification to the atmospheric correction of SeaWiFS ocean color images over turbid waters. *Cont. Shelf Res.* 25 (4), 539–555.
- Lee, Z.-P., Ahn, Y.-H., Mobley, C., Arnone, R., 2010. Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform. *Opt. Express* 18, 26313–26342.
- Lee, Z., Carder, K.L., Mobley, C.D., Steward, R.G., Patch, J.S., 1999. Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Appl. Opt.* 38, 3831–3843.
- Lee, Z., Shang, S., Lin, G., Chen, J., Doxaran, D., 2016. On the modeling of hyperspectral remote-sensing reflectance of high-sediment-load waters in the visible to shortwave-infrared domain. *Appl. Opt.* 55, 1738–1750.
- Min, K., Cuiffi, J.D., Mathers, R.T., 2020. Ranking environmental degradation trends of plastic marine debris based on physical properties and molecular structure. *Nat. Commun.* 11, 1–11.
- Moolayil, J., Moolayil, J., John, S., 2019. *Learn Keras for deep neural networks*. Springer.
- Mouw, C.B., Greb, S., Aurin, D., DiGiorgio, P.M., Lee, Z., Twardowski, M., Binding, C., Hu, C., Ma, R., Moore, T., Moses, W., Craig, S.E., 2015. Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sens. Environ.* 160, 15–30.
- Mueller, J.L., Davis, C., Arnone, R., Frouin, R., Carder, K., Lee, Z., Steward, R., Hooker, S., Mobley, C.D., McLean, S., 2000. Above-water radiance and remote sensing reflectance measurements and analysis protocols. *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation Revision 2*, 98–107.
- Murray, N.J., Keith, D.A., Bland, L.M., Ferrari, R., Lyons, M.B., Lucas, R., Pettorelli, N., Nicholson, E., 2018. The role of satellite remote sensing in structured ecosystem risk assessments. *Sci. Total Environ.* 619, 249–257.
- Pahlevan, N., Roger, J.C., Ahmad, Z., 2017. Revisiting short-wave-infrared (SWIR) bands for atmospheric correction in coastal waters. *Opt. Express* 25, 6015–6035.
- Pan, Y., Shen, F., Verhoef, W., 2017. An improved spectral optimization algorithm for atmospheric correction over turbid coastal waters: A case study from the Changjiang (Yangtze) estuary and the adjacent coast. *Remote Sens. Environ.* 191, 197–214.
- Pettorelli, N., Schulte to Bühne, H., Tulloch, A., Dubois, G., Macinnis-Ng, C., Queirós, A. M., Keith, D.A., Wegmann, M., Schrodt, F., Stellmes, M., Sonnenschein, R., Geller, G. N., Roy, S., Somers, B., Murray, N., Bland, L., Geizendorfer, I., Kerr, J.T., Broszeit, S., Leitão, P.J., Duncan, C., El Serafy, G., He, K.S., Blanchard, J.L., Lucas, R., Mairota, P., Webb, T.J., Nicholson, E., Rowcliffe, M., Disney, M., 2018. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. *Remote Sens. Ecol. Conserv.* 4 (2), 71–93.
- Preisendorfer, R.W., 1976. *Hydrologic optics vol. 1: Introduction*. Springfield: National Technical Information Service. Also available on CD, Office of Naval Research.
- Ruddick, K.G., Ovidio, F., Rijkeboer, M., 2000. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl. Opt.* 39, 897–912.
- Schroeder, T.H., Behnert, I., Schaale, M., Fischer, J., Doerffer, R., 2007. Atmospheric correction algorithm for MERIS above case-2 waters. *Int. J. Remote Sens.* 28 (7), 1469–1486.
- Shang, S., Lee, Z., Lin, G., Li, Y., Li, X., 2019. Progressive scheme for blending empirical ocean color retrievals of absorption coefficient and chlorophyll concentration from open oceans to highly turbid waters. *Appl. Opt.* 58, 3359–3369.
- Shehhi, M.R., Gherboudj, I., Zhao, J., Ghedira, H., 2017. Improved atmospheric correction and chlorophyll-a remote sensing models for turbid waters in a dusty environment. *ISPRS J. Photogramm. Remote Sens.* 133, 46–60.
- Shi, W., Wang, M., 2007. Detection of turbid waters and absorbing aerosols for the MODIS ocean color data processing. *Remote Sens. Environ.* 110 (2), 149–161.
- Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. *Opt. Express* 19, 9783–9800.
- Stumpf, R.P., Arnone, R., Gould, R.W., Martinolich, P., Ransibrahmanakul, V. (Eds.), 2002. A Partially-Coupled Ocean-Atmosphere Model for Retrieval of Water-Leaving Radiance from SeaWiFS in Coastal Waters. NASA Tech. Memo, Greenbelt, MD.
- Thuillier, G., Hersé, M., Simon, P., Mandel, H., Gillotay, D., 1998. Observation of the solar spectral irradiance from 200 nm to 870 nm during the ATLAS 1 and ATLAS 2 missions by the SOLSPEC spectrometer. *Metrologia* 35, 689.
- Vanhellemont, Q., 2020. Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre-and decametre-scale satellite imagery using autonomous hyperspectral radiometry. *Opt. Express* 28, 29948–29965.
- Vanhellemont, Q., Ruddick, K., 2015. Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8. *Remote Sens. Environ.* 145, 89–106.
- Wang, M., 2002. The Rayleigh lookup tables for the SeaWiFS data processing: accounting for the effects of ocean surface roughness. *Int. J. Remote Sens.* 23 (13), 2693–2702.
- Wang, M., 2005. A refinement for the Rayleigh radiance computation with variation of the atmospheric pressure. *Int. J. Remote Sens.* 26 (24), 5651–5663.
- Wang, M., 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: simulations. *Appl. Opt.* 46, 1535–1547.
- Wang, M., Bailey, S., 2001. Correction of the sun glint contamination on the SeaWiFS ocean and atmosphere products. *Appl. Opt.* 40, 4790–4798.
- Wang, J., Lee, Z., Wang, D., Shang, S., Wei, J., Gilerson, A., 2021a. Atmospheric correction over coastal waters with aerosol properties constrained by multi-pixel observations. *Remote Sens. Environ.* 265, 112633.
- Wang, Y., Lee, Z., Wei, J., Shang, S., Wang, M., Lai, W., 2021b. Extending satellite ocean color remote sensing to the near-blue ultraviolet bands. *Remote Sens. Environ.* 253, 112228.
- Wang, M., Shi, W., 2007. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing. *Opt. Express* 15, 15722–15733.
- Wang, M., Tang, J., Shi, W., 2007. MODIS-derived ocean color products along the China east coastal region. *Geophys. Res. Lett.* 34 (6).
- Werdell, P.J., Franz, B.A., Lefler, J.T., Robinson, W.D., Boss, E., 2013. Retrieving marine inherent optical properties from satellites using temperature and salinity-dependent backscattering by seawater. *Opt. Express* 21, 32611–32622.
- Xue, C., Chen, S., Lee, Z., Hu, L., Shi, X., Lin, M., Liu, J., Ma, C., Song, Q., Zhang, T., 2021. Iterative near-infrared atmospheric correction scheme for global coastal waters. *ISPRS J. Photogramm. Remote Sens.* 179, 92–107.
- Yu, X., Lee, Z., Shen, F., Wang, M., Wei, J., Jiang, L., Shang, Z., 2019. An empirical algorithm to seamlessly retrieve the concentration of suspended particulate matter from water color across ocean to turbid river mouths. *Remote Sens. Environ.* 235, 111491.
- Zibordi, G., Mélin, F., Berthon, J.-F., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Vandemark, D., Feng, H., Schuster, G., Fabbri, B.E., Kaitala, S., Seppälä, J., 2009.

- AERONET-OC: a network for the validation of ocean color primary products. *J. Atmos. Oceanic Technol.* 26 (8), 1634–1651.
- Zibordi, G., Holben, B.N., Talone, M., D'Alimonte, D., Slutsker, I., Giles, D.M., Sorokin, M.G., 2021. Advances in the ocean color component of the Aerosol Robotic Network (AERONET-OC). *J. Atmos. Oceanic Technol.* 38 (4), 725–746.
- Zou, S., Xu, W., Zhang, R., Tang, J., Chen, Y., Zhang, G., 2011. Occurrence and distribution of antibiotics in coastal water of the Bohai Bay, China: impacts of river discharge and aquaculture activities. *Environ. Pollut.* 159 (10), 2913–2920.