Journal of Remote Sensing
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

Atmospheric Correction of Coastal Waters Based on Satellite–AERONET-OC Matchups

Xu Li, Zhongping Lee*, Jinyan Xie, Tianhao Wang, Lufei Zheng, Daosheng Wang, Xiaolong Yu, and Shaoling Shang

State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China.

*Address correspondence to: ZhongPing.Lee@umb.edu

Citation: Li X, Lee Z, Xie J, Wang T, Zheng L, Wang D, Yu X, Shang S. Atmospheric Correction of Coastal Waters Based on Satellite–AERONET-OC Matchups. *J. Remote Sens.* 2025;5:Article 0886. https://doi.org/10.34133/remotesensing.0886

Submitted 18 March 2025
Revised 11 August 2025
Accepted 25 August 2025
Published 16 October 2025

Atmospheric correction (AC) is a critical step in ocean color remote sensing, particularly for coastal waters that still face challenges from high concentrations of suspended materials and absorbing aerosols. To address these limitations, this study presents a novel AC algorithm, termed ACA-SIM (atmospheric correction based on satellite–in situ matchup data), based on extensive matchups between satellite measurements (ρ_t) and in situ remote sensing reflectance (R_{rs}) from Aerosol Robotic Network-Ocean Color (AERONET-OC), with neural networks as the processing tool. Unlike the Ocean Color-Simultaneous Marine and Aerosol Retrieval Tool (OC-SMART algorithm, which employs a similar strategy but relies on simulated data, ACA-SIM uses real-world matchups between ρ_t and in situ R_{rs}, capturing sensor-specific effects such as striping and encompassing a wide range of water and aerosol properties. Validations with independent AERONET-OC measurements demonstrated that ACA-SIM outperformed both the NASA Standard and OC-SMART, achieving higher accuracy of R_{rs} across all spectral bands. In particular, for the blue bands, ACA-SIM reduced the mean absolute percentage difference (MAPD) of derived R_{rs} to ~15%, compared to an MAPD of ~32% by OC-SMART, and maintained robust performance even under challenging conditions. Moreover, when applied to Moderate-Resolution Imaging Spectroradiometer-Aqua images in highly turbid and dust- or smoke-affected regions, such as the Bohai and Yellow Seas, the West Coast of North Africa, and areas impacted by Australian bushfires, ACA-SIM demonstrated exceptional capability in minimizing striping effects and generating reliable R_{rs} products. This study advances AC techniques for coastal waters and reinforces the importance of the AERONET-OC network. Furthermore, it lays a foundation for extending ACA-SIM to other satellite sensors, enabling the generation of consistent and accurate ocean color products across multiple satellite platforms.

Introduction

Coastal waters are vital ecosystems, playing a crucial role in biogeochemical cycles and representing some of the most dynamic and complex regions in the marine environment. Approximately 70% of the world's largest cities, with populations exceeding one million, are within 100 km of coastlines. In recent decades, these waters have experienced increasing occurrences of red tides and harmful algal blooms, making them a key area of interest for ocean research [1,2]. Systematic and sustained observations of the rapidly changing coastal ecosystem are thus an essential component of the Earth observing system [3]. Compared to traditional in situ measurements, large-scale and high-frequency satellite remote sensing has become indispensable for such observations [4]. It is thus not surprising to see that, since the launch of the Coastal Zone Color Scanner (CZCS), especially the Moderate-Resolution Imaging Spectroradiometer (MODIS), over 20 years of valuable continuous data of the global ocean have been provided via ocean-color satellites. However, accurately retrieving water properties in coastal waters from satellite measurements remains challenging [5]. High concentrations of dissolved and suspended materials in water and complex aerosol properties in such regions complicate achieving accurate retrievals of water properties from ocean-color satellite data.

Atmospheric correction (AC) is a crucial step in processing satellite ocean color measurements. Its goal is to remove the effects of atmosphere and sea surface reflectance from the total radiance signal (L_t) measured by a satellite sensor at the top of the atmosphere (TOA), intending to accurately extract the signal of water body, i.e., water-leaving radiance (L_w), or remote sensing reflectance (R_{rs}) that is converted from L_w [6]. This process provides the foundation for subsequent estimations of bio-optical parameters from R_{rs} [7,8]. AC algorithms focus on estimating the aerosol contributions in atmospheric contributions, which introduces the dominant sources of uncertainties in the final bio-optical products [5].

The classical approach for removing aerosol contributions relies on the "black pixel" assumption in the near-infrared (NIR) bands [6,9], which assumes that L_w in the NIR bands is negligible due to the strong absorption by water. Such an approach has shown great success for open ocean waters, which

has been a standard for processing data from the earliest CZCS to modern ocean-color satellites such as MODIS and the Visible Infrared Imaging Radiometer Suite (VIIRS). However, for turbid coastal waters, the water contribution is no longer negligible in the NIR bands, leading to an overestimation of the aerosol contribution and an underestimation of L_w, which can even result in negative values for L_w in the blue bands [10–13]. In recent decades, considerable effort has been invested in developing AC algorithms for coastal waters, resulting in a wide range of algorithms developed (see the "Brief Description of 2 Typical and Contrasting AC Approaches for Coastal Waters" section for more details). These include approaches modeling L_w in the NIR bands [14–17], or extending the black pixel assumption to the shortwave infrared (SWIR) bands [18–20], or incorporating the blue band in the process of AC [21], or using the darkest pixel within an image [22,23], each showing various levels of success.

In parallel, since neural networks (NNs) or machine learning (ML) is showing superior nonlinear problem-solving capabilities each day, schemes based on NN or ML have also been developed for AC problems in oceanic and coastal waters [24–30]. Such a scheme relies on high-quality and representative data to train the model [31], where synthetic datasets were commonly used [26,31–33]. Such simulations, however, generally ignore the land-adjacent effects in coastal regions (e.g., [34]), and may not be able to account for all natural variations in aerosols and water's bio-optical properties in coastal and open-ocean regions. More importantly, the various specifics of sensors, including sensor noises such as striping effects, were not incorporated. Consequently, the simulated TOA radiance may not precisely match that measured by a satellite sensor for the same suite of atmosphere and ocean properties.

Ideally, it would be best to use accurate and inclusive data from both field measurements and satellite L_t to train data-driven (DD) algorithms, either NN or ML. Unfortunately, due to the high cost of conducting ship surveys, the matchups between field-measured high-quality R_{rs} and satellite-measured L_t are very limited. Therefore, few studies have developed DD-based AC algorithms using matchup in situ and satellite datasets. On the other hand, the Aerosol Robotic Network-Ocean Color (ANOC hereafter) maintains more than 40 stations globally, offering long-term observation data covering a wide range of water types that include blue coastal waters to CDOM (colored dissolved organic matter)-dominated green–yellow waters and highly turbid brown waters, as well as various atmospheric conditions [35].

Conventionally, the ANOC-R_{rs} products have been used as an important source to validate satellite R_{rs} products [36–39]. In this study, to further demonstrate the value and usefulness of ANOC-R_{rs}, we developed an AC algorithm, termed ACA-SIM (atmospheric correction based on satellite–in situ matchup data), based on satellite and ANOC matchups, explicitly targeting coastal waters and using NNs as the processing tool. The key feature of ACA-SIM is a compiled matchup dataset between ANOC-R_{rs} from the various sites and satellite measurements. This dataset provides extensive coverage of challenging conditions, including large solar zenith angles (SOZA), sensor-specific effects, stray light, absorbing aerosols, cloud shadows, and adjacency effects from nearby land, factors that were often overlooked in previous similar approaches. Using L_t data from MODIS-Aqua as an example, ACA-SIM was developed for processing MODIS-Aqua measurements, and its performance was independently validated using data from field measurements and evaluated using images of MODIS-Aqua. The results highlight the effectiveness of ACA-SIM in processing coastal waters, even under dust and smoke. This approach has strong potential for further extending to other sensors and establishes a benchmark for generating consistent R_{rs} products in coastal waters across different satellites.

This article is organized as follows: the "Brief Description of 2 Typical and Contrasting AC Approaches for Coastal Waters" section provides a brief overview of the theoretical basis of AC and 2 typical and contrasting AC approaches (NASA Standard and Ocean Color-Simultaneous Marine and Aerosol Retrieval Tool [OC-SMART]) used for processing satellite ocean-color images. The "AC Based on Satellite–In Situ Matchup Data" section introduces the development of ACA-SIM. In the "In Situ Data for Validation" section, in situ measurements used for validation are presented. The "Results and Evaluations" section presents the performance of ACA-SIM and compares it with the NASA Standard and OC-SMART, along with demonstrations with MODIS-Aqua images. The "Discussion" section further discusses reasons for the better performance of ACA-SIM under challenging conditions, such as absorbing aerosols and striping. Finally, the "Conclusion" section provides a summary of this effort.

Materials and Methods

Brief description of 2 typical and contrasting AC approaches for coastal waters

Overall relationship

The total radiance (L_t) measured by a satellite sensor at the TOA can be generally expressed as:

$$L_t(\lambda) = L_r(\lambda) + L_a(\lambda) + L_x(\lambda) + t(\lambda)L_w(\lambda). \quad (1)$$

Here, L_r represents contributions from Rayleigh scattering, and L_a denotes contributions from aerosols (including those of Rayleigh-aerosol multiple scatterings). Note that the sum of L_r and L_a makes up ~80%, or more (especially in the NIR-SWIR domain), of L_t [5]. L_x here represents the noises arising from sun glint and whitecaps or foams. t is the diffuse transmittance of the atmosphere. The desired property is the water-leaving radiance (L_w), as it contains information about water constituents. Here, for brevity, the effect of gases in the atmosphere is omitted in Eq. 1. Since radiance is also a function of solar irradiance, reflectance is generally used, and the TOA reflectance (ρ_t) is defined as:

$$\rho_t(\lambda) = \frac{\pi L_t(\lambda)}{F_0 \cos(\theta_s)}. \quad (2)$$

F_0 is the extraterrestrial solar irradiance [40], and θ_s is the solar zenith angle. Therefore, Eq. 1 can be written as

$$\rho_t(\lambda) = \rho_r(\lambda) + \rho_a(\lambda) + \rho_x(\lambda) + t_s(\lambda)t(\lambda)\pi R_{rs}(\lambda) \quad (3)$$

Here, $\rho_{r,a,x}$ are the reflectance due to Rayleigh, aerosol, and surface effects, respectively, and t_s is the diffuse transmittance of the solar radiation. $R_{rs}(\lambda)$ is the remote-sensing reflectance of a water body, which is defined as:

$$R_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d^{0+}(\lambda)}, \quad (4)$$

with $E_d^{0+}(\lambda)$ being the downwelling irradiance just above the surface. The objective of an AC algorithm is to obtain, as accurately as possible, $R_{rs}(\lambda)$ from $\rho_t(\lambda)$. As described in detail in IOCCG Report #10 [5], a wide range of AC algorithms have been developed in the past decades. Given the limited space of this article and the scope of this study, we highlight here 2 commonly used but contrasting AC algorithms: the NASA Standard and OC-SMART. We further compare the performance of ACA-SIM with that of the 2 algorithms.

NASA Standard approach

In Eq. 3, ρ_r can be exactly calculated with information on sun-sensor geometry and at sea level atmospheric pressure [41–43], with ρ_x estimated based on geometry and wind speed [44–46]. The overall challenge lies in calculating ρ_a. This is achieved with the "black pixel" assumption that R_{rs} is 0 in the NIR or SWIR bands; therefore, ρ_t data in these bands after removing ρ_r and ρ_x become ρ_a, which is used to estimate aerosol characteristics and then extrapolated to the blue–green bands, subsequently to obtain R_{rs} of these bands [18,41,47,48]. For waters where R_{rs} in the NIR is no longer negligible, algorithms were developed to estimate R_{rs}(NIR) [14–17], or to extend the NIR bands to SWIR [18,47], in order to overcome the limitations of this assumption.

This classical AC approach is incorporated into the SeaWiFS Data Analysis System (SeaDAS) endorsed by the National Aeronautics and Space Administration (NASA) and is the most widely used AC algorithm. In this study, the default setting of SeaDASv8.4, which employs the iterative NIR scheme, is referred to as the NASA Standard.

OC-SMART

As early as the 1990s, it was demonstrated that R_{rs} could be estimated from L_t or ρ_t using NNs [26], a much simpler, more efficient, but implicit approach compared to the NASA Standard AC system. This NN-based system was later improved [27,28] for better accuracy in coastal waters by integrating a coupled radiance transfer system for forward modeling. More recently, for AC of multiple sensors with the same NN architecture, also with more accurate data representing L_t of various oceanic and atmospheric conditions, Fan et al. [32,33] proposed the OC-SMART AC algorithm, with data used for training the AC algorithm by large simulations of the coupled atmosphere–ocean system, where more complex aerosol conditions were introduced [49]. Specially, the simulations considered multiple scattering effects and interactions between atmospheric and oceanic components, enabling more accurate simulations of L_t over a wide range of aerosol optical depth (AOD) and R_{rs} across diverse water properties [8]. Finally, after training, the algorithm uses Rayleigh-corrected radiance and sun-sensor geometries as input to estimate R_{rs} of various satellite sensors. It is necessary to note that such simulations cannot account for sensor-specific (e.g., striping) and land-adjacency effects. In addition, the spectral range used for training is 412 to 869 nm, which may encounter similar issues to the NASA Standard for highly turbid waters, where L_w is no longer negligible in the NIR domain. For this effort, we applied the latest available version (Python_Linux_v2.1) of the OC-SMART code provided by Fan et al. [33], which is available from the open-source repository (the Light and Life website, http://www.rtatmocn.com/oc-smart/).

AC based on satellite–in situ matchup data

Both the NASA Standard and OC-SMART, as well as many other AC algorithms, have demonstrated various successes in processing satellite ocean-color measurements [5,50]; however, challenges remain. These include observations made under large solar zenith angles, strong sunglint, strong absorbing aerosols, and cloud shadows, to name a few. In particular, these approaches rely on perfect L_t data, where the sensor's noises or striping effects were not included in the development of the AC algorithms [32,33], which were then propagated to the final bio-optical products [51,52]. A separate and dedicated step is required to remove such sensor noises in these AC schemes [53]. On the other hand, there are rich and high-quality R_{rs} data available since the installation of the first AERONET-OC site in 2002; therefore, there is a potential to develop a DD AC algorithm through matching up data between satellite ρ_t and ANOC-R_{rs}, termed ACA-SIM, where sensor effects in ρ_t could be accounted for, although implicitly.

Overview of the strategy

The crucial factor to DD algorithms lies in the quality and representativeness of the dataset. Thus, the overall strategy for the development of ACA-SIM is as follows: (a) compile a large, high-quality, and representative dataset that includes both satellite ρ_t and ANOC-R_{rs}, and (b) train a robust NN system for the estimation of R_{rs} from satellite ρ_t. In essence, ACA-SIM follows a similar strategy to OC-SMART, but its notable distinction is the use of satellite measurements along with in situ R_{rs} data. Compared to synthetic datasets based on RT simulations, in situ measurements offer a more authentic representation of R_{rs} spectra. Furthermore, the matchup between satellite data and in situ R_{rs} covers a wide range of atmospheric conditions, observing geometries, and preserves the noise present in real satellite measurements. Numerical simulations that follow radiative transfer, however, are difficult to accurately model challenging observation conditions, such as high SOZA, strong sunglint, and sensor-specific effects, such as striping and straylight. These factors often lead to a decrease in the accuracy of AC algorithms, or even distortions, especially when applied to images.

R_{rs} from AERONET-OC

AERONET-OC measurements

All AERONET-OC data used in this effort were downloaded from the AERONET website (http://aeronet.gsfc.nasa.gov), where multi-band normalized water-leaving radiance (L_{wn}) and auxiliary information are included. There are 3 levels of version-3 AERONET-OC data: Level 1.0 (minimized platform disturbance), Level 1.5 (cloud and wave anomaly removal), and Level 2.0 (fully quality-assured normalized water-leaving radiance) [35,54]. In this study, Level 1.5 data, rather than Level 2.0, were utilized because Level 1.5 contains up to 210,000 spectra from January 2002 to March 2024, compared to approximately 80,000 spectra under Level 2.0. As always, extensive and inclusive data are a key factor for DD-based algorithms. It should be noted that Level 1.5 data in version 3 may contain potential spectral inconsistencies. However, the application of the band-shifting procedure effectively minimizes such effects and filters out unreasonable spectra (see the "Band-shifting of AERONET-OC to match that of MODIS" section and Appendix A).

The L_{wn} data acquired from AERONET-OC sites were converted to R_{rs} as follows:

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

AERONET-OC has utilized CE-318 and CE-318T automated sun photometers, installed on offshore platforms such as lighthouses, oceanographic towers, and oil rigs worldwide, to measure optical data of water and atmosphere (see Fig. 1) [55]. Data from AERONET-OC have several advantages: (a) measurements are taken using a unified system and protocol; (b) calibration is performed using the same reference source and methods; and (c) data processing is conducted using uniform code.

Although it is the same type of instrument, the settings of the instrument (especially the spectral bands) at each AERONET-OC site are not necessarily the same [35]. Therefore, it is necessary to do a band shift to convert the multi-band $R_{rs}^{AO}(\lambda)$ obtained at each site to R_{rs} of the MODIS bands. A brief description of this band-shift approach is provided below, with details included in Appendix A.

Band-shifting of AERONET-OC to match that of MODIS

To shift the spectral bands of AERONET-OC, which may vary from site to site, to match that of MODIS-Aqua, we first constructed hyperspectral R_{rs} from multi-band R_{rs}^{AO}, and then obtained R_{rs} at the MODIS bands through spectral convolution. For expanding multi-band R_{rs} to hyperspectral R_{rs}, the hyperspectral optimization processing exemplar (HOPE) developed by Lee et al. [56] was used. Finally, after quality control, 168,572 MODIS-wavelength R_{rs} spectra were obtained from these AERONET-OC measurements after the construction of hyperspectral R_{rs}. As examples, Fig. 2 shows hyperspectral and MODIS-band R_{rs} spectra from 4 representative sites, indicating that these R_{rs} covered waters from blue clear waters to green water with high concentrations of colored, dissolved organic matter and highly turbid brownish water. The other AERONET-OC sites supplement the gaps among these 4 sites with varying water constituents and concentrations.

MODIS-Aqua data for the development of ACA-SIM

In this study, data from MODIS-Aqua were chosen for the development and demonstration of ACA-SIM. This is because, from 2002 to 2024, MODIS-Aqua offers the longest coverage and the largest dataset to match that with AERONET-OC measurements. Level 1A data of MODIS-Aqua (L_t) were obtained from NASA's Ocean Color Website (https://oceancolor.gsfc.nasa.gov/). It is necessary to first generate the MODIS-Aqua L1B and geo files before exporting the required Level 2 products. Standard R_{rs}, ρ_t, 4 angles, and atmospheric parameters were simultaneously obtained from the l2gen module in SeaDAS. NASA Standard R_{rs} were obtained using the default settings (multi-scattering aerosol model and NIR correction with up to 10 iterations). Additionally, the established criteria L2-FLAGS for MODIS-Aqua are used to ensure data quality. A variety of measurement conditions were encountered for these coastal sites, which include HILT (observed radiance very high or saturated), LAND (pixel is over land), CLDICE (probable cloud or ice contamination), STRAYLIGHT (probable stray light contamination), and MODGLINT (moderate sun glint contamination).

Matchup criteria

AERONET-OC sites encompass a wide variety of water types, which ensures the applicability of ACA-SIM across diverse coastal waters and atmospheric conditions. The following criteria were taken to match ANOC-R_{rs} with ρ_t from MODIS-Aqua:

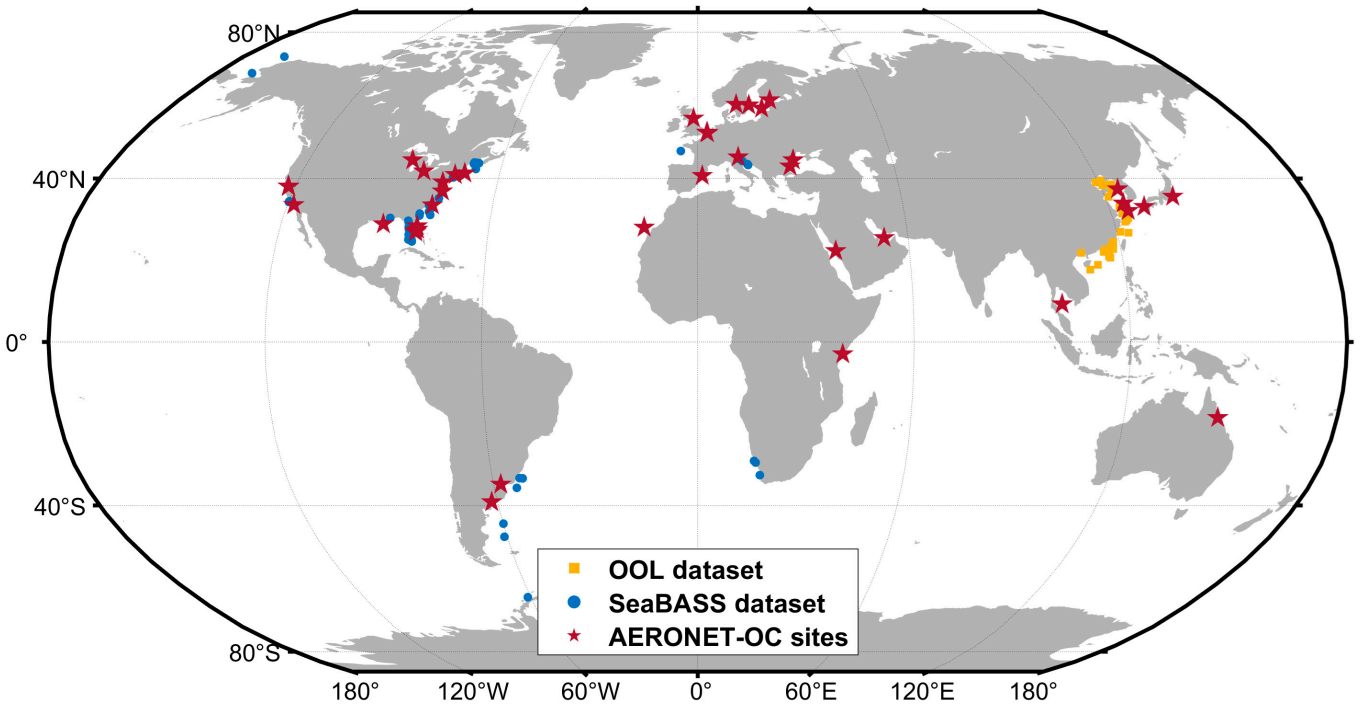

Fig. 1. The spatial distribution of field-measured data used in this study. Red pentagrams, blue circles, and yellow squares represent data from the AERONET-OC sites, Optical Oceanography Lab (OOL, Xiamen University, China), and SeaBASS, respectively. Note that field measurements with water depth > 1,000 m were excluded from this study.

1. Satellite data where the water signal was completely obscured by anomalies (e.g., land, clouds, and very high or saturated radiance) were excluded, while pixels moderately affected by contamination (e.g., stray light, moderate sunglint, and cloud shadows) were retained; thus, the developed ACA-SIM could handle such complex conditions.
2. The time window is restricted to ±1 h to take account of the more rapid variability of coastal waters, rather than the common ±2 h [23,29,36,57] used for the study of oceanic waters. If multiple ANOC-R_{rs} spectra matched the same MODIS-Aqua data, only the one with the smallest time difference is retained.
3. Since the focus is coastal waters where the spatial heterogeneity of water properties is much higher than that of oceanic waters, the nearest pixel of MODIS-Aqua matching an AERONET-OC site was used, rather than the conventional average of 3×3 pixels [36,57] around a site.

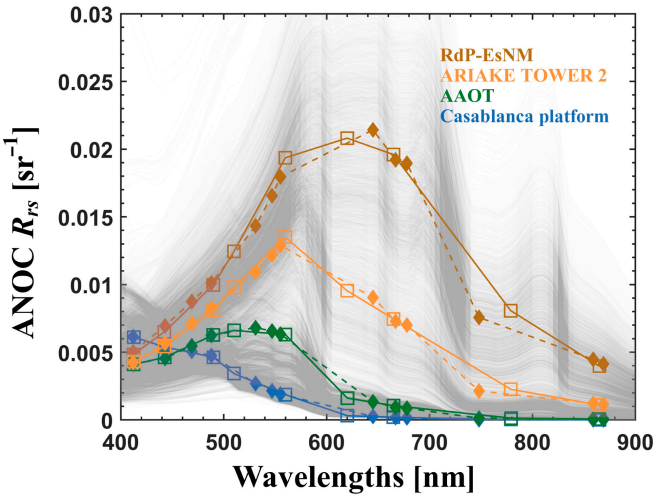

Fig. 2. R_{rs} spectra of 4 representative sites with different water types (Casablanca_Platform, $n = 5{,}968$; AAOT, $n = 4{,}958$; ARIAKE_TOWER_2, $n = 2{,}821$; RdP-EsNM, $n = 4{,}500$). Solid line–square represents the average spectra of the original multi-band ANOC-R_{rs}, and dashed curve–diamond denotes R_{rs} band-shifted to MODIS-Aqua bands through HOPE. Gray curves with transparency indicate all HOPE-reconstructed hyperspectral R_{rs} spectra for the 4 sites.

The above resulted in 8,818 matchups between MODIS-Aqua ρ_t and ANOC-R_{rs} for the training of ACA-SIM. Table 1 and Fig. 3 present the range and distributions of key parameters of this dataset, which largely determine the model's applicability. For the sun-sensor angles, the SOZA ranges from 9° to 77.8°, which is wider than the common upper limit of 70°

Table 1. Parameters and ranges of the ACA-SIM training dataset

Parameter	Minimum value	Maximum value
SOZA (°)	9.0	77.8
SOAA (°)	−179.9	179.9
SEZA (°)	0.03	65.4
SEAA (°)	−177.1	178.4
Chla (mg/m^3)	0.02	41.4
AOD(412)	0.01	0.77

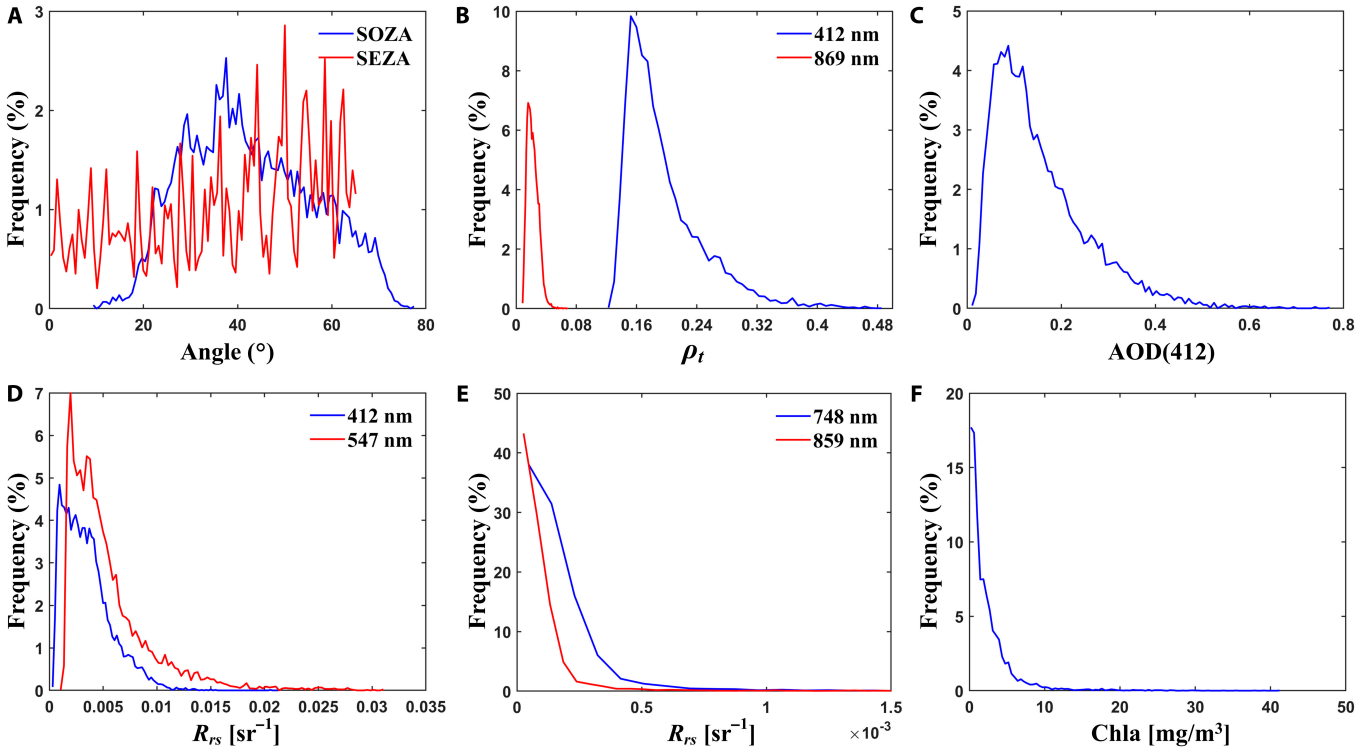

Fig. 3. Histogram of the data in the training dataset. (A) Solar zenith angle (blue) and sensor zenith angle (red); (B) ρ_t(412) (blue) and ρ_t(869) (red); (C) AOD(412); (D) R_{rs}(412) (blue) and R_{rs}(547) (red); (E) R_{rs}(748) (blue) and R_{rs}(859) (red); (F) Chla.

[29,32], indicating that ACA-SIM may perform better in high-latitude oceans and during morning or evening [25]. The sensor zenith angle (SEZA) ranges from 0.03° to 65.4°. The solar azimuth angle (SOAA) and sensor azimuth angle (SEAA) range from −179.9° to 179.9° and from −177.1° to 178.4°, respectively, covering nearly the entire possible range of satellite observation. For the water parameters, in addition to the wide range of R_{rs} values (see Fig. 3D and E), the estimated chlorophyll-a concentration (Chla) based on ANOC-R_{rs} ranges from 0.02 to 41.4 mg/m^3, indicating highly dynamic water environments. Here, the Chla values were taken from the associated AERONET-OC dataset, with the estimation algorithms detailed in Zibordi et al. [35]. Regarding atmospheric conditions, aerosol optical density at 412 nm [AOD(412)] ranges from 0.01 to 0.77, indicating widely varying aerosol loadings.

Multilayer perceptron

Multilayer perceptron (MLP) is an artificial neural network consisting of an input layer, one or more hidden layers, and an output layer. Each layer contains multiple neurons fully connected to the neurons in adjacent layers. The MLP uses a feedforward structure, where information flows from the input layer through the hidden layers to the output layer. Each neuron applies an activation function, which introduces nonlinearity and enables the network to learn complex patterns in the data. MLPs are trained through backpropagation, where the network iteratively adjusts its weights to minimize the error between predicted and actual outputs [58]. In this study, after many rounds of tries, it was found that in an MLP architecture with 4 hidden layers (512, 256, 128, and 64 neurons), using the rectified linear unit activation function [59,60] between the hidden layers worked well. Also, a sigmoid activation function at the output layer was used to prevent negative values. The selection of a perfect NN architecture remains elusive. We experimented with different configurations to determine the optimal number of hidden layers and neurons and found that deep networks yielded better performance [61]. Training was stopped based on early stopping criteria, monitoring validation loss to prevent overfitting [62].

The input to the NN consists of 14 bands of ρ_t at wavelengths 412, 443, 469, 488, 531, 547, 555, 645, 667, 678, 748, 859, 869, and 1,240 nm, along with sun-sensor angles, and atmospheric parameters (water vapor, pressure, ozone, NO_2 components, and relative humidity). All these parameters can be obtained through SeaDAS. The inclusion of 1,240 nm is important for processing measurements over highly turbid waters. However, because some detectors at the 1,640-nm band are broken [63], we excluded this band to ensure the quality of the dataset. We also omitted the SWIR bands due to their low signal-to-noise ratio. The output layer consists of 13 bands of R_{rs} corresponding to the visible and NIR wavelengths from 412 to 869 nm. Figure 4 summarizes the input, output, and hidden layers of this model.

In situ data for validation

In addition to using the AERONET-OC data, a large dataset collected in various coastal environments was used to evaluate ACA-SIM and the 2 aforementioned contrasting AC algorithms. These in situ R_{rs} were obtained from the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) and the Optical Oceanography Laboratory (OOL). Note that since this study focuses on AC of coastal waters, the in situ data were limited to water depths of less than 1,000 m. Data with a bottom depth greater than 1,000 m were excluded.

The SeaBASS is a centralized repository maintained by NASA's Ocean Biology Processing Group (OBPG) for in situ oceanographic and atmospheric data, supporting the validation of satellite data products and algorithm development [64,65]. The dataset is accessible via the SeaBASS website (https://seabass.gsfc.nasa.gov) and has undergone various quality control procedures to ensure data reliability. It is necessary to note that the R_{rs} data deposited in SeaBASS, which are at the MODIS bands, were provided by colleagues from the international community; thus, the measurement approaches and processing methods may vary from group to group.

Spanning from June 2003 to December 2023, the OOL collected a wide range of in situ R_{rs} spectra during cruises in the coastal waters of China (see Fig. 1). The in situ dataset comprises measurements obtained using both the skylight-blocked approach (SBA) [66,67] and the above-water approach (AWA) [68,69]. Following the NASA protocol [68], the AWA employs a handheld spectroradiometer to measure the upwelling total

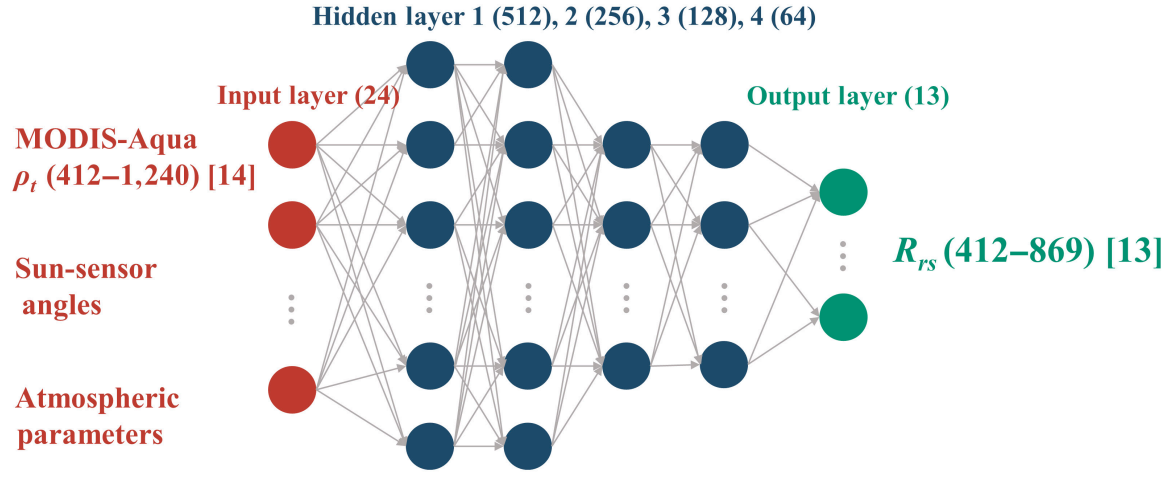

Fig. 4. The structure of ACA-SIM.

radiance, downwelling sky radiance, and graycard-reflected radiance, with R_{rs} finally calculated from these measurements [70]. Based on the SBA concept, water-leaving radiance and downwelling irradiance right above the surface were measured simultaneously. Subsequently, R_{rs} spectra were obtained after shading error correction [67,71]. The radiometers used for these measurements have a spectral range of 320 to 950 nm with a spectral resolution of ~3 nm, and the resulting hyperspectral R_{rs} were converted to MODIS bands via spectral convolution. More details about these field measurements can be found in Dong et al. [72] and Shang et al. [73].

Following the same matchup procedure between MODIS-Aqua and AERONET-OC, 159 pairs of spectra were obtained between MODIS-Aqua and these field measurements.

Results

Methods of accuracy evaluation

The performance of AC algorithms was comprehensively evaluated using the following statistical metrics, which are coefficient of determination (R^2), mean absolute percentage difference (MAPD), mean absolute difference (MAD), and bias, defined as follows"

$$\mathrm{MAPD} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{y_i}\right| \times 100\%, \tag{6}$$

$$\mathrm{MAD} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_i\right|, \tag{7}$$

$$\mathrm{Bias} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - y_i\right), \tag{8}$$

where x_i and y_i are AC algorithms derived and known (in situ) values of R_{rs}, respectively.

Performance of ACA-SIM with the training data

The performance of ACA-SIM was first evaluated using the training dataset, where 20% of the data were reserved as a validation set to assess the model's effectiveness, while 80% of the data were used for training ACA-SIM. Figure 5 compares R_{rs} obtained from ACA-SIM with ANOC for spectral bands from 412 to 869 nm of the validation set. The evaluation results of the training dataset are presented in Appendix B. For the algorithm inter-comparison, the R_{rs} products obtained from the NASA Standard and OC-SMART were also compared to the same ANOC-R_{rs} for spectral bands in this domain (see Figs. 6 and 7).

Overall, ACA-SIM demonstrates excellent performance across all bands among the 3 algorithms; in particular, there are substantial improvements in the blue and red bands, where the NASA Standard frequently produced negative values (see Fig. 6). On statistical measures, for the blue–green bands, ACA-SIM achieves a coefficient of determination (R^2) consistently above 0.89, with no negative values at the 412- and 443-nm bands, and the MAPD reduced to ~15% for the 3 blue bands (412, 443, and 488 nm). For R_{rs} in blue–green (412 to 555 nm), the NASA Standard achieved R^2 in a range of 0.54 to 0.93, while MAPD is in a range of ~15% to 54% (see Fig. 6), with the green bands (531, 547, and 555 nm) showing the best performance. This is due to the fact that coastal waters have the strongest

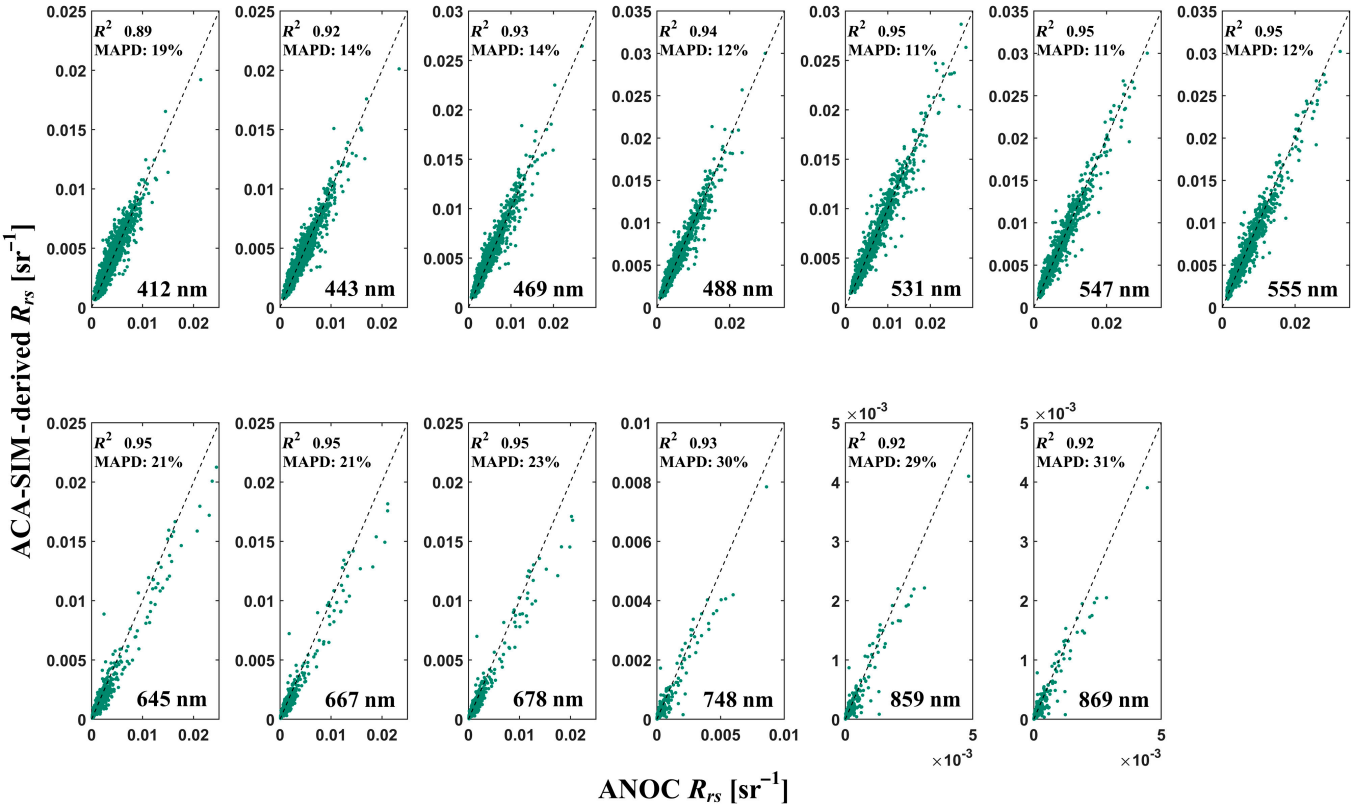

Fig. 5. Validation of R_{rs} derived from ACA-SIM using the validation dataset (n = 1,764).

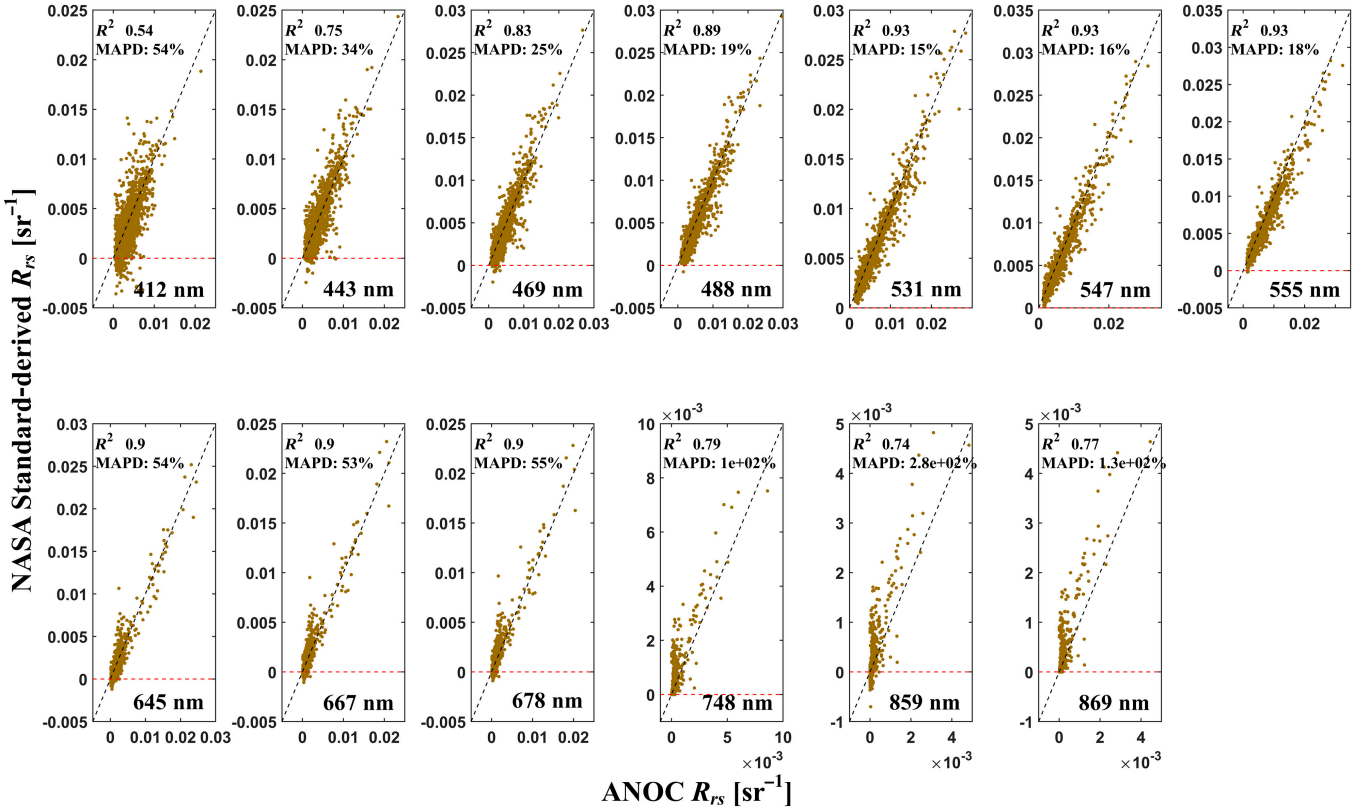

Fig. 6. Validation of R_{rs} derived from the NASA Standard using the validation dataset ($n = 1{,}764$).

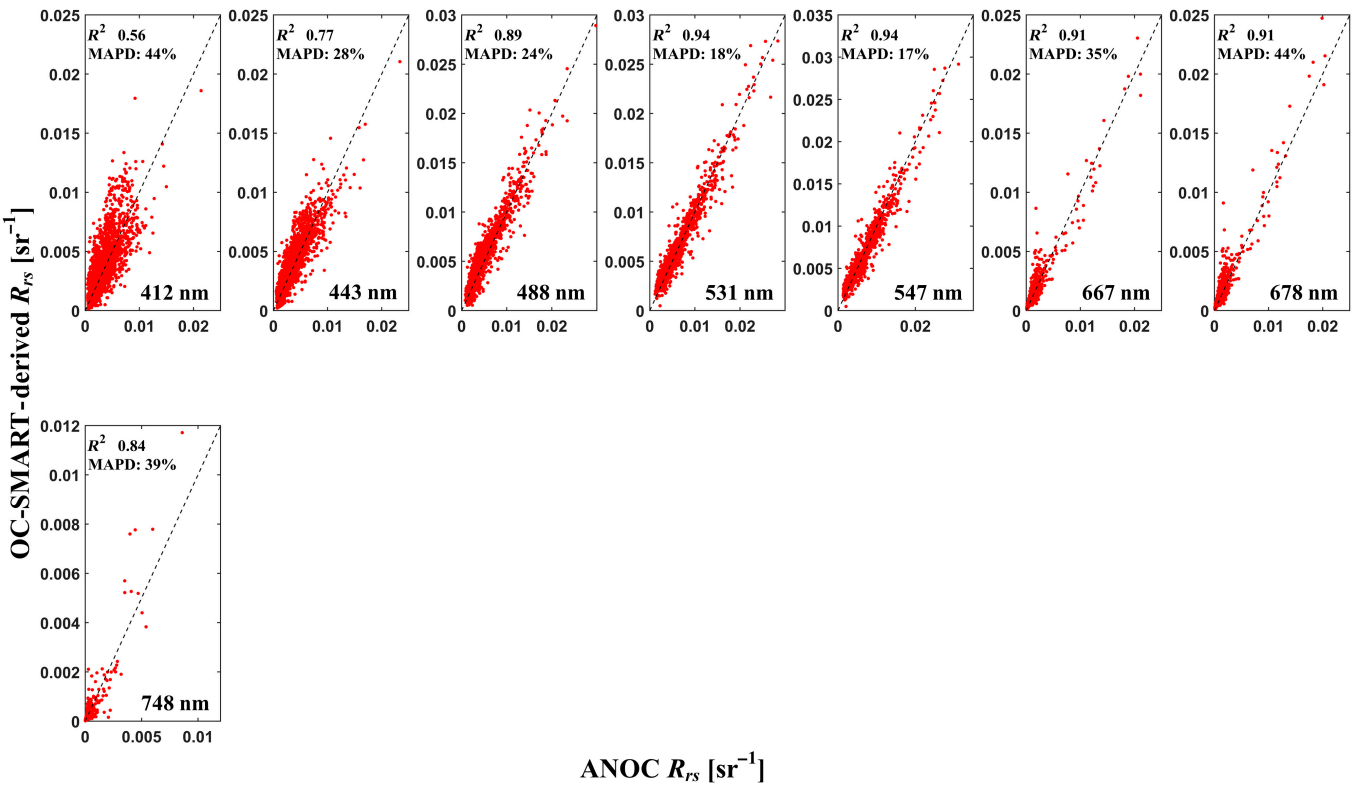

Fig. 7. Validation of R_{rs} derived from OC-SMART using the validation dataset ($n = 1{,}764$).

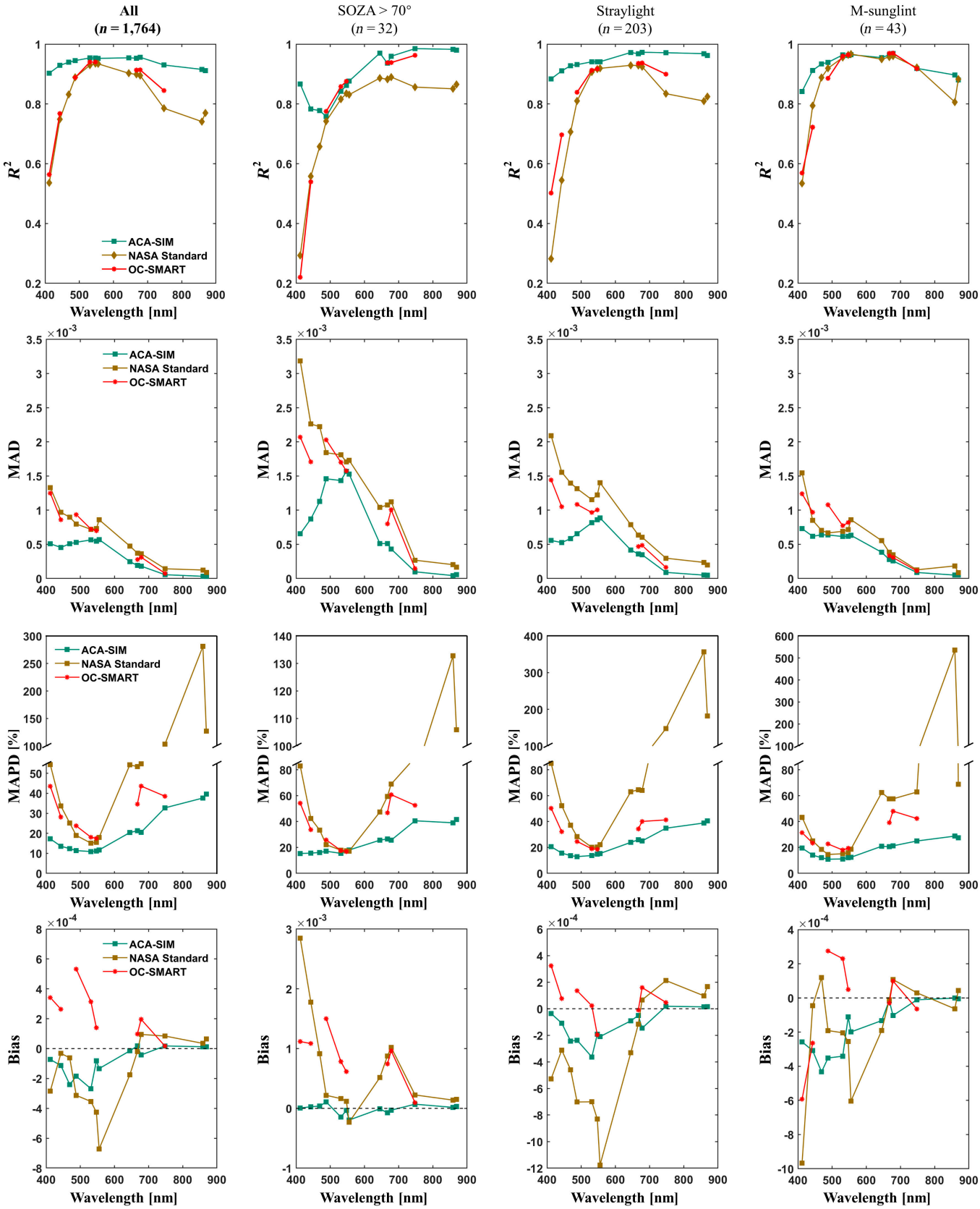

Fig. 8. Statistical results of the NASA Standard-, OC-SMART-, and ACA-SIM-derived R_{rs} compared with ANOC-R_{rs} of the validation set. The validation set is further categorized into high solar zenith angles (>70°), moderate sunglint, and stray light.

Table 2. The AERNET-OC sites used in leave-one-site-out cross-validation

ID	Site	*n*	ID	Site	*n*
S1	AAOT	953	S10	LISCO	422
S2	Casablanca_Platform	662	S11	Lucinda	339
S3	Chesapeake_Bay	101	S12	MVCO	639
S4	Galata_Platform	461	S13	Section-7_platform	287
S5	Gloria	455	S14	Socheongcho	119
S6	Gustav_Dalen_Tower	440	S15	Thornton_C-power	121
S7	Helsinki_Lighthouse	239	S16	USC_SEAPRISM	661
S8	Irbe_Lighthouse	153	S17	Venise	1,637
S9	Kemigawa_Offshore	105	S18	WaveCIS_Site_CSI_6	421

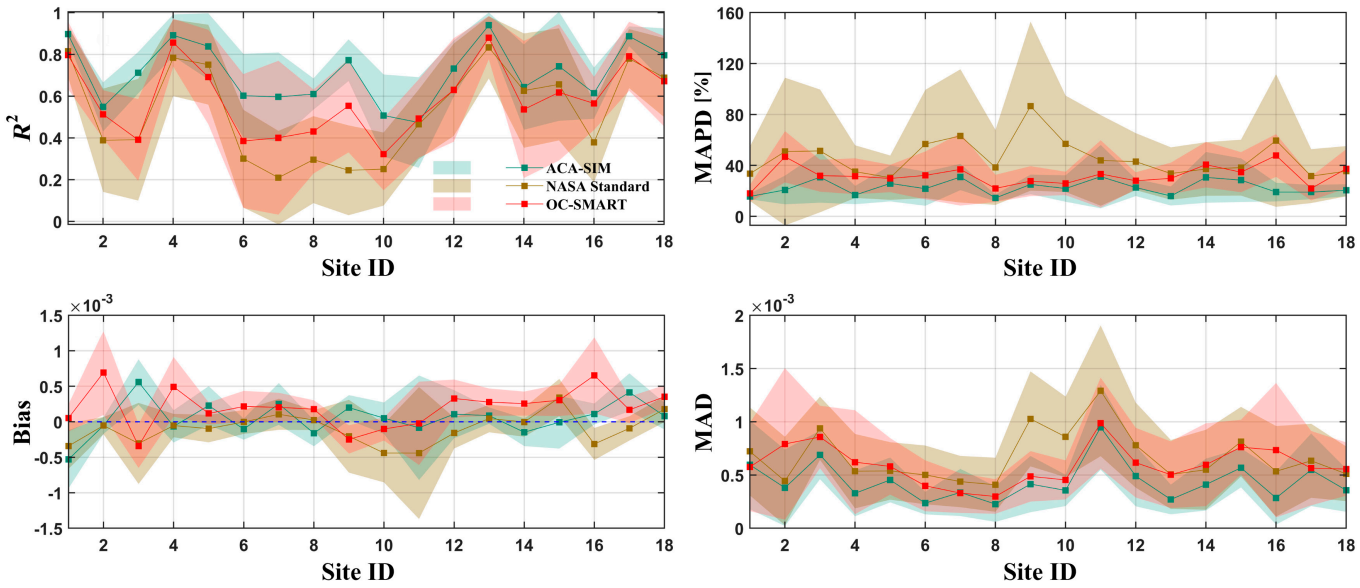

Fig. 9. Validation at AERONET-OC sites using leave-one-site-out validation.

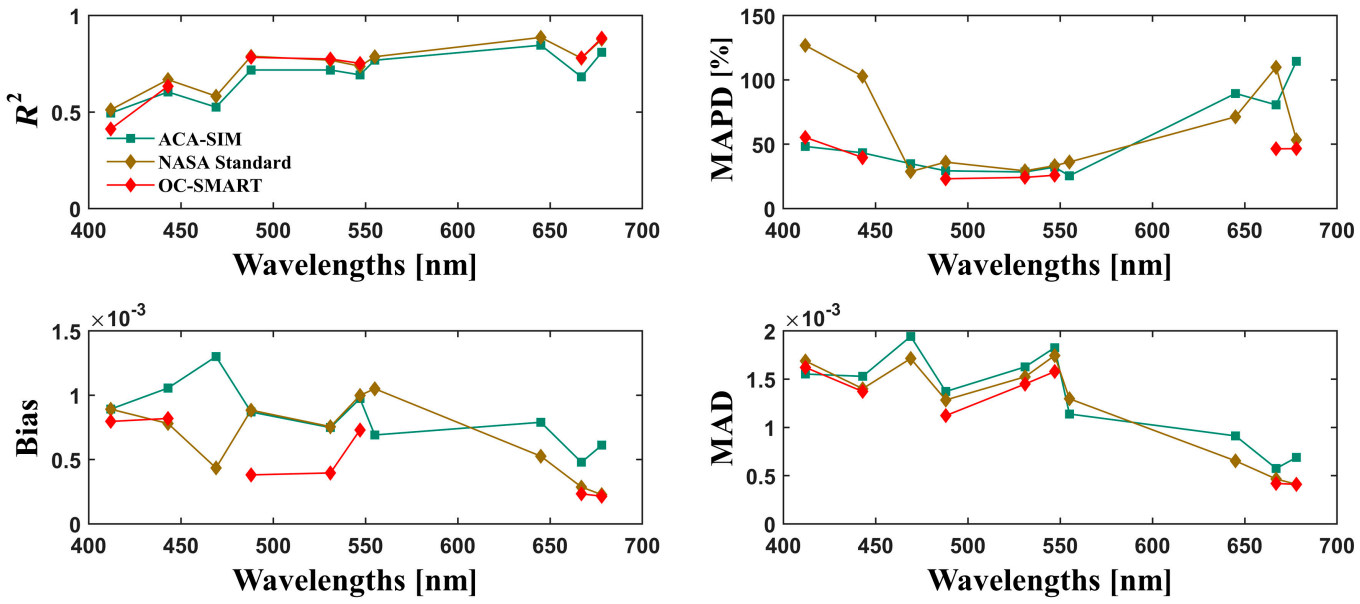

Fig. 10. Statistical results of ACA-SIM, OC-SMART, and NASA Standard of the in situ datasets.

signal in the green bands (see Fig. 2), thus less affected by noises due to measurements and AC. OC-SMART does not produce negative R_{rs} either (see Fig. 7), and the R^2 values ranged from 0.56 to 0.94, and MAPD ranged from ~17% to 44% (~32% for the 3 blue bands), with the green bands (531 and 547 nm) also showing the best performance compared to other bands.

For the red–NIR bands (645- to 869-nm range), where R_{rs} values are generally very low due to strong absorption by water molecules and the ANOC-R_{rs} might include unquantifiable uncertainties, ACA-SIM also obtained considerably better R_{rs} compared to the NASA Standard (see Figs. 5 to 7). For the NASA Standard, the MAPD values range from ~54% to 280%, while the MAPD of ACA-SIM is ~20 to 40%. OC-SMART did not produce R_{rs} for the 859- and 869-nm bands, and the achieved results were comparable to ACA-SIM at 748 nm, with MAPD and R^2 values of 39% and 0.84, respectively.

Figure 8 summarizes the statistical results from the 3 AC algorithms presented above. Additionally, it highlights the breakdown results for observations under high solar zenith angle (>70°), moderate sun glint (L2_FLAGS_20 = 1), and stray light conditions (L2_FLAGS_8 = 1), which were indicated by the flags defined in SeaDAS. It was found that the R_{rs} in the blue bands obtained from the NASA Standard and OC-SMART were strongly affected by these observation conditions, although

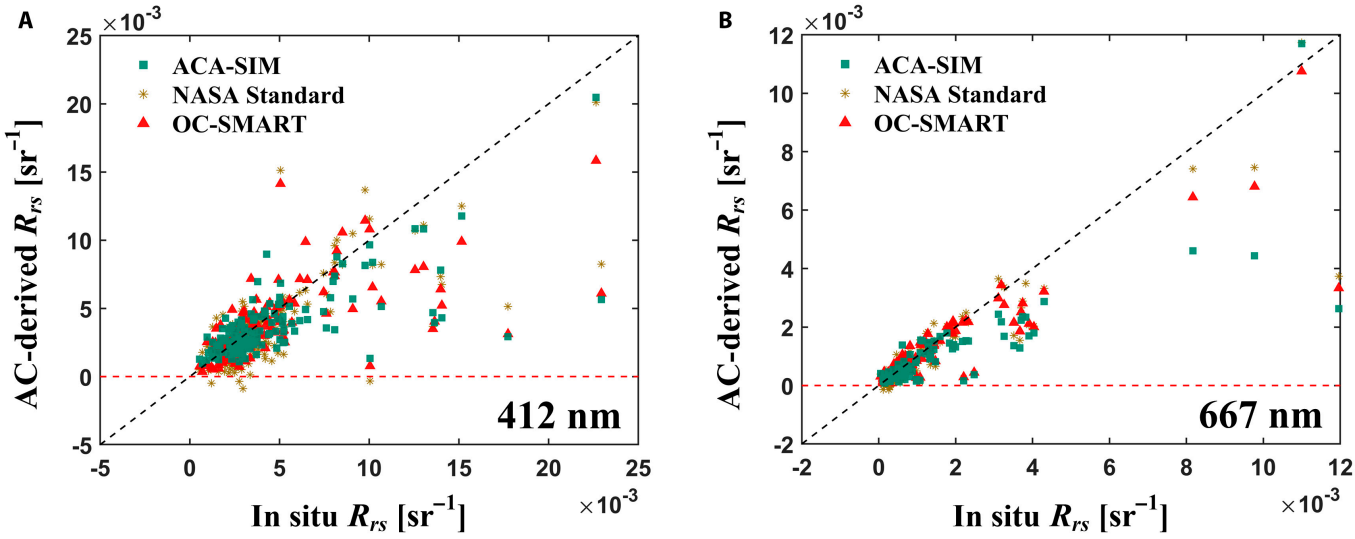

Fig. 11. Scatterplots between in situ R_{rs} and that derived by the NASA Standard, OC-SMART, and ACA-SIM from MODSI-Aqua measurements (n = 147 and 106 for 412 and 667 nm, respectively): (A) 412 nm; (B) 667 nm.

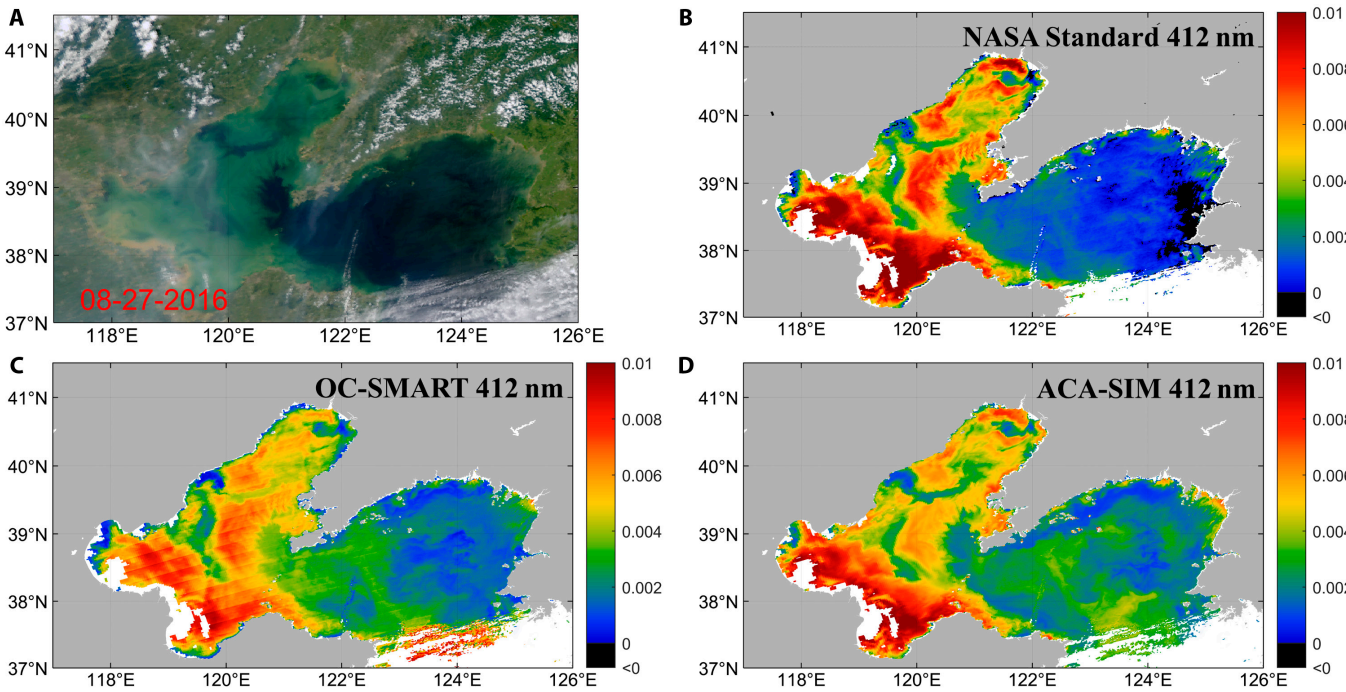

Fig. 12. Results of R_{rs}(412) by the NASA Standard, OC-SMART, and ACA-SIM over the Bohai and Yellow Seas collected by MODIS-Aqua on 2016 August 27. (A) RGB true color image. R_{rs}(412) obtained from the NASA Standard (B), OC-SMART (C), and ACA-SIM (D).

reasonable results were obtained for the green bands. However, ACA-SIM produced substantially improved R_{rs} at the 412- and 443-nm bands, as well as more consistent R_{rs} with ANOC-R_{rs} for bands from green to NIR for the challenging observation conditions. Overall, the above comparisons indicate a strong performance (higher R^2, lower MAD and MAPD, and much smaller bias) of ACA-SIM for these coastal waters.

Evaluation of ACA-SIM with independent data

Validation via leave-one-site-out

Since DD-based algorithms are data-dependent, a big concern is whether such algorithms can be applied to data not included in the training. To evaluate the concept and applicability of ACA-SIM, a leave-one-site-out validation method was employed for an independent test and validation. Specifically, among the n AERONET-OC sites, data from the ith site S_i is excluded from the dataset for algorithm development and simply used for validation, while data from the remaining sites S_j ($j = 1, \ldots, n; j \neq i$) are used for model training. Since data from S_i is entirely out of the training process, this constitutes an independent validation. By iterating this process across all n sites (resulting in n ACA-SIM$_i$), a series of independent validations were obtained. For this evaluation, it is important to use statistically significant sample sizes (we here kept a minimum of 100 data points); therefore, in total, 18 sites (see Table 2) were selected for this leave-one-site-out validation.

Figure 9 shows the performance of ACA-SIM across the 18 sites, along with that of the NASA Standard and OC-SMART. Since it is an iterative process over the 18 sites, there are 18 numbers of R^2, MAPD, MAD, and bias. Note that these values represent the averages of the spectral bands of produced R_{rs}, with wavelengths matching those of OC-SMART (see Fig. 7 for wavelengths). As examples, results of representative blue–green-red bands of 412, 547, and 667 nm are included in Appendix C. Overall, it is found that no matter which site is left out for the development evaluation, the R^2 values of ACA-SIM at all sites (0.71 ± 0.14) are better than those of the NASA Standard (0.53 ± 0.21) and OC-SMART (0.58 ± 0.21), while the MAPD and MAD values are lower than those of the NASA Standard and OC-SMART, with a small bias. The consistently better performance across all sites indicates that even when one site is excluded, the remaining sites still encompass a diverse set of water and atmosphere properties, allowing the model to effectively learn and generalize across different conditions. This further highlights that the present AERONET-OC data used for developing ACA-SIM possess a sufficiently broad dynamic

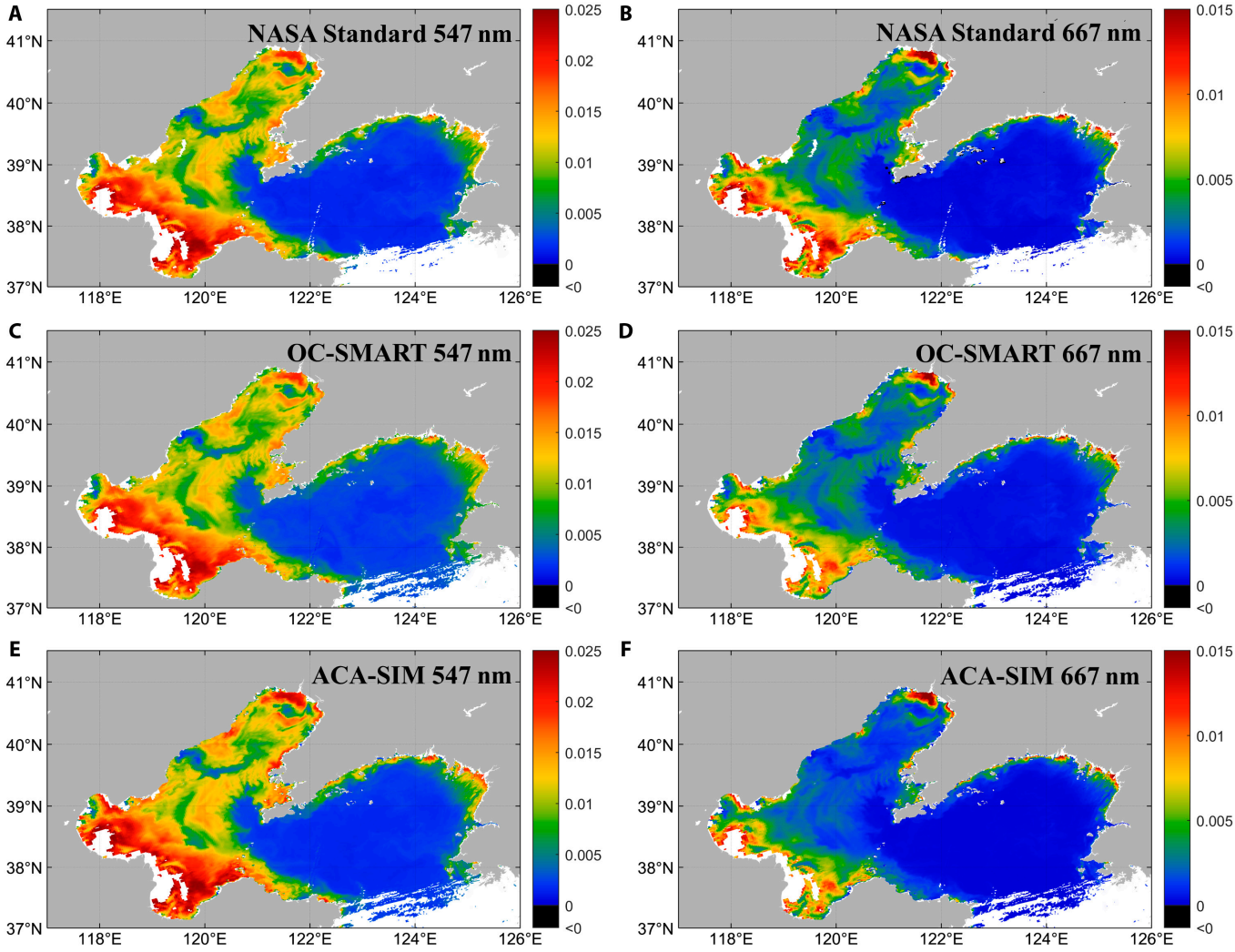

Fig. 13. Results of R_{rs}(547) and R_{rs}(667) by the NASA Standard, OC-SMART, and ACA-SIM for the same scene and date as Fig. 12. (A, C, and E) R_{rs}(547) obtained from the NASA Standard, OC-SMART, and ACA-SIM, respectively; (B, D, and F) R_{rs}(667) obtained from the NASA Standard, OC-SMART, and ACA-SIM, respectively.

range and representativeness, although further adding sites would certainly be helpful. Thus, it is confident to apply ACA-SIM after it is trained with all available data.

Evaluation using ship-based measurements

As a common practice [15,22,32], we further evaluated the performance of ACA-SIM using the compiled ship-based in situ data, with Fig. 10 showing the performance measures. Statistically, there is no clear ranking in the performance of the 3 algorithms, as the NASA Standard shows slightly better R^2 values for blue to red bands, while ACA-SIM and OC-SMART obtain better measures in MAPD for these bands, and the MAD values are similar for the 3 algorithms. It is necessary to keep in mind that it is very difficult to maintain the same quality of R_{rs} from different groups when it was measured in situ, as it involves diverse cruise-specific measurement conditions, the use of different sensors, different measurement geometries, differences in operator's handling and experience, and different processing procedures [64,66,70]. Thus, a large portion of the difference between MODIS-Aqua R_{rs} (no matter which AC algorithm) and in situ R_{rs} is very likely due to different qualities of the measured R_{rs} as well as the mismatches in space and time between satellite and field data. AERONET-OC data, however, through standardization in sensors, measurement geometry, and data processing, can remove many of the inconsistencies in obtaining field R_{rs}.

To get more insights into this comparison, Fig. 11 shows scatterplots of R_{rs} at the blue and red bands; again, no clear advantages among the 3 AC algorithms, except that both OC-SMART and ACA-SIM avoided the negative values occasionally produced by the NASA Standard. For R_{rs}(412), the MAPD of OC-SMART and ACA-SIM is reduced by ~60% compared to that of the NASA Standard (see Fig. 10). On the higher end of R_{rs} values, it appears that all 3 ACs exhibited an

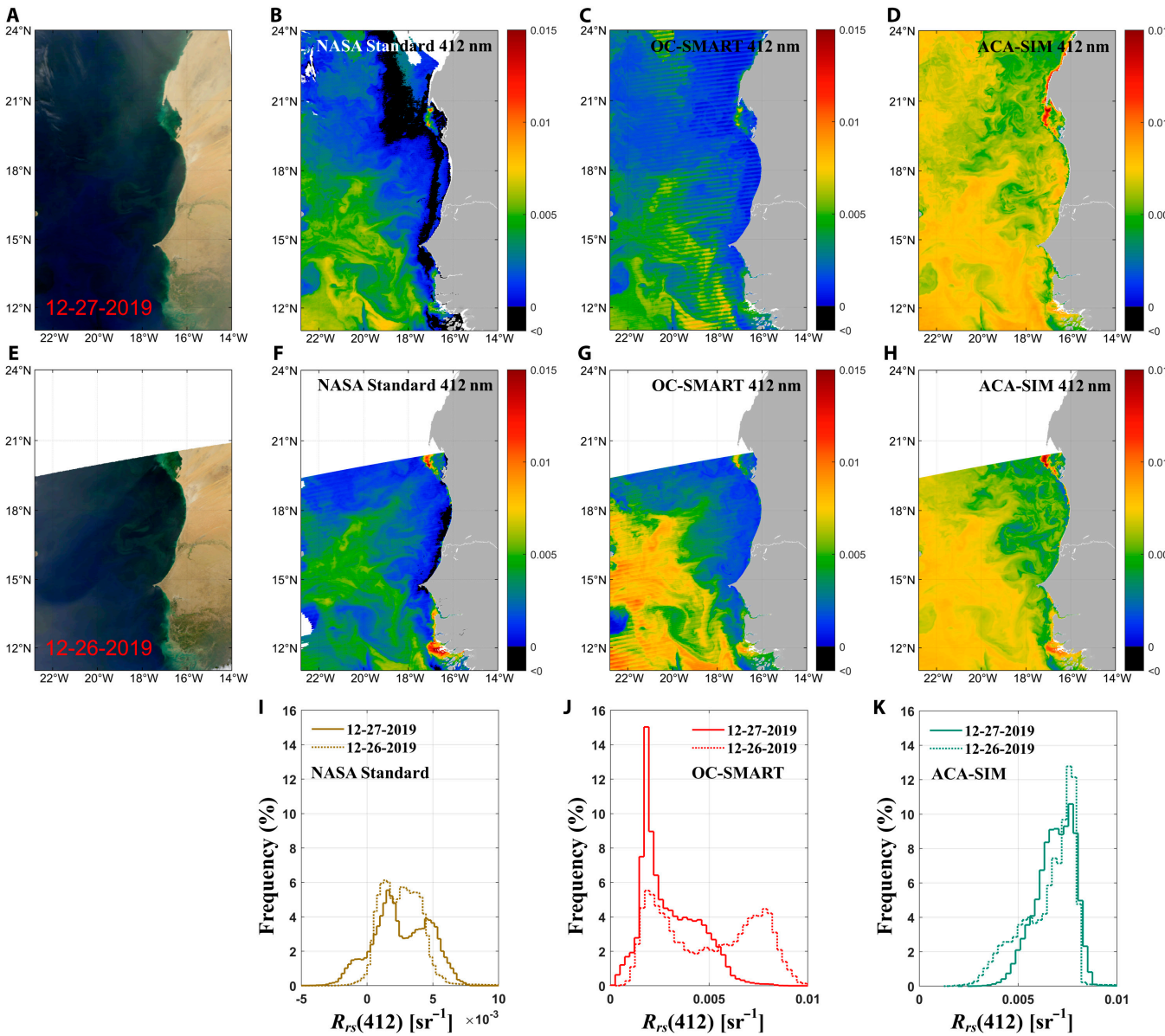

Fig. 14. Application to the MODIS-Aqua images in the West Coast of North Africa. (A to D) RGB composite image and R_{rs} values at 412 nm derived from the NASA Standard, OC-SMART, and ACA-SIM on 2019 December 27, respectively; (E to H) Same as (A) to (D), but for 2019 December 26. (I to K) Histograms of the R_{rs}(412) values of the 2 days.

"underestimation". This may further highlight the challenges in obtaining consistent, high-quality R_{rs} from diverse groups in field measurements.

Application to MODIS-Aqua images

The ultimate goal and value of an algorithm are in its applicability to process satellite ocean color images; we thus further evaluated the performance of ACA-SIM through its application to satellite images of various coastal environments, especially challenging scenarios, such as highly turbid waters or highly absorbing aerosols. It is important to note that these scenarios often occur simultaneously and are coupled. In addition, we also present the results of the NASA Standard and OC-SMART for comparison.

Bohai and Yellow Seas

The Bohai and Yellow Seas are semi-enclosed by industrial cities of China and the Korean peninsula, experiencing frequent and strong air pollution. Additionally, the Yellow River estuary in the Bohai Sea carries a substantial sediment load into the ocean, resulting in highly turbid waters. Figure 12A shows an RGB composite image of 2016 August 27, collected by MODIS-Aqua, where the distinct sediment-laden waters of the Bohai Sea contrast with the relatively clear waters of the Yellow Sea. For this MODIS-Aqua measurement, the NASA Standard obtained negative $R_{rs}(412)$ along the eastern coast of the Yellow Sea (see Fig. 12B), a region characterized by the coupling of turbid waters and absorbing aerosols [74,75]. There are also discrete negative $R_{rs}(412)$ along the northwestern coast of the Bohai Sea. Both ACA-SIM and OC-SMART, on the other hand, did not produce negative $R_{rs}(412)$ and obtained overall higher $R_{rs}(412)$ values than the NASA Standard, effectively capturing the spatial variations of the water masses in the Bohai and Yellow Seas. In extremely turbid coastal waters, OC-SMART $R_{rs}(412)$ values are lower than those derived by ACA-SIM, consistent with the underestimation observed by Song et al. [29] in the Bohai and Yellow Seas. In particular, in the southeastern corner of the image, where sparse southwest-oriented cloud streaks are present, $R_{rs}(412)$ from OC-SMART exhibits extremely high anomalies, likely due to the influence of cloud edges and/or shadows. This phenomenon is also commonly observed in the results from the NASA Standard; therefore, the products are usually masked out. However, thanks to the use of real data for training, ACA-SIM demonstrates exceptional resistance to such interference, producing results highly consistent with the surrounding water masses. These results suggest that ACA-SIM is more effective in handling noise in real image data, enhancing the overall accuracy and consistency of the retrieved ocean color products.

For the 547- and 667-nm bands, the 3 AC methods exhibit similar R_{rs} values and spatial distributions in the Bohai and Yellow Seas (see Fig. 13). This further confirms that, compared to the 412-nm band, the influence of turbid waters and absorbing

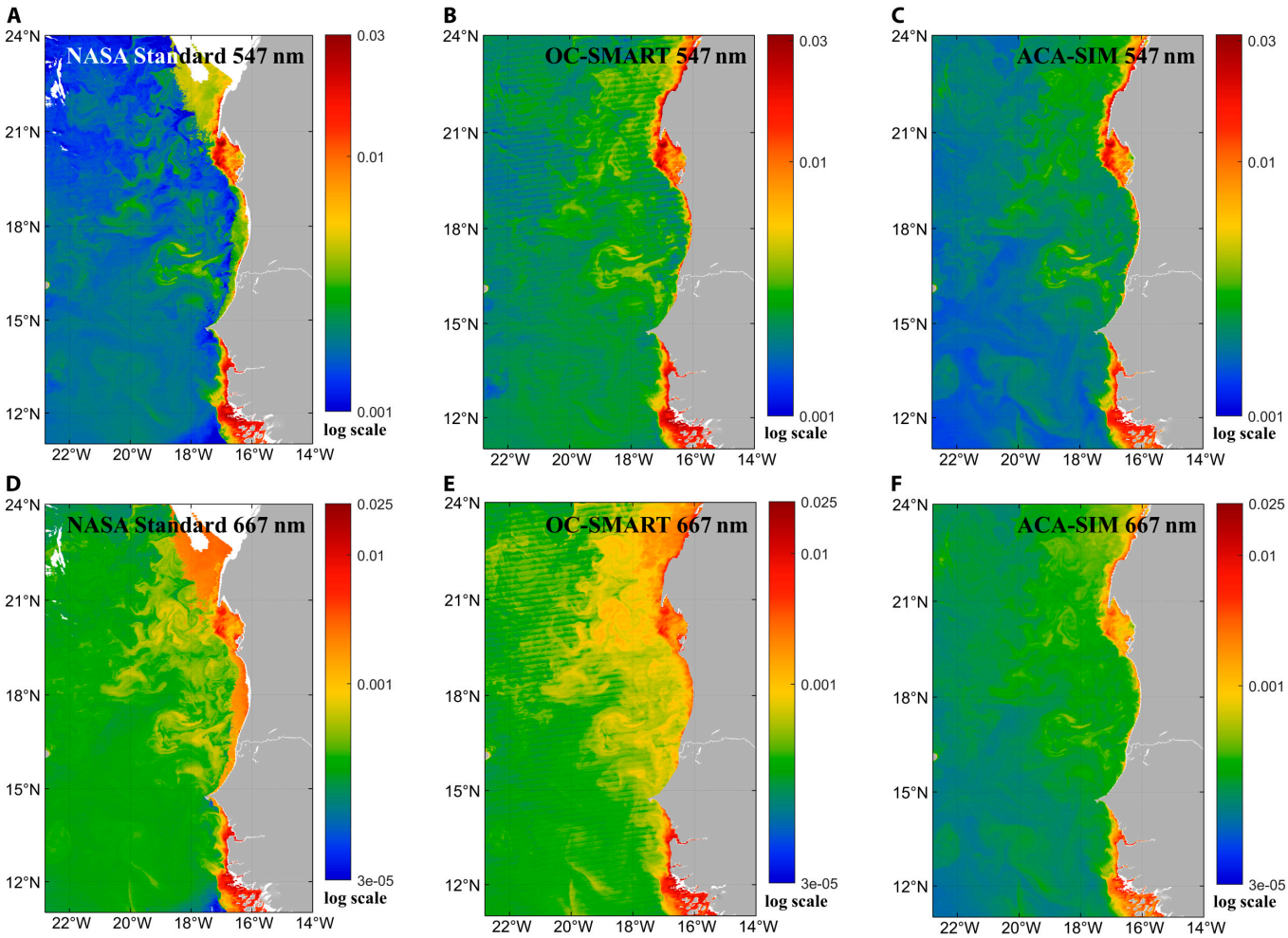

Fig. 15. Similar to Fig. 14, but showing the results at 547 and 667 nm (A to F) using a logarithmic scale.

aerosols on the AC is less pronounced at these wavelengths. On the other hand, the high consistency in R_{rs} among the 3 independently developed AC approaches provides mutual validation of their effectiveness and accuracy. However, ACA-SIM demonstrates a notable advantage in minimizing the striping effects (more details in the "Striping artifacts" section), which remain visible in the results from both the NASA Standard and OC-SMART (see Fig. 12).

West Coast of North Africa

Driven by the Northeast trade winds, vast amounts of dust from the Sahara Desert are transported over the waters of the West Coast of North Africa, extending as far as the Cape Verde Islands in the North Atlantic and forming one of the largest airborne dust events over the ocean (see Fig. 14A). As a representative type of strongly absorbing aerosols, dust aerosols frequently cause failures in AC algorithms [76–78]. In the MODIS-Aqua RGB image of 2019 December 27, a thick yellow dust aerosol layer blankets the coastal waters, with the highest concentration observed in the northern region. On the other hand, deep green turbid waters are visible in the northern bays and the southernmost river delta. For these MODIS-Aqua measurements, the NASA Standard shows a band of negative $R_{rs}(412)$ parallel to the coastline (see Fig. 14B), along with widespread negative values in the northern region where dust aerosol loads were high. Beyond the negative $R_{rs}(412)$ values, the low $R_{rs}(412)$ regions (dark blue) in the NASA Standard results align with the distribution of the dust aerosol layer, suggesting potential underestimation in these areas.

OC-SMART successfully avoided negative $R_{rs}(412)$ values. However, the resulting $R_{rs}(412)$ is as low as ~0.0018 sr^{-1} for far-offshore Atlantic waters, indicating dust effects in the R_{rs} results. ACA-SIM, however, avoided negative $R_{rs}(412)$, and the resulting $R_{rs}(412)$ is around 0.008 sr^{-1} for offshore waters, a value more consistent with oceanic waters. To further check the validity of the $R_{rs}(412)$ values, MODIS-Aqua measurements from 1 day earlier (2019 December 26) were also processed and presented (Fig. 14E to H), a day with different sky conditions (Fig. 14A vs. Fig. 14E). Note that due to satellite orbital variations, the same region is located at the edge of the image on 2019 December 26, extending only up to approximately 20°N. It is found that the $R_{rs}(412)$ values on 2019 December 26 are markedly (~53% on average) higher than those of 2019 December 27, when processed by OC-SMART. However, the difference of $R_{rs}(412)$ between the 2 days on average is just ~7% when obtained from ACA-SIM. Figure 14I to K further highlight the distribution of $R_{rs}(412)$ of the 2 consecutive days obtained from the 3 ACAs, where there are large changes in $R_{rs}(412)$ from the NASA Standard and OC-SMART, but it remains nearly the same for ACA-SIM. As water's properties do not change drastically day by day, these results indicate that the $R_{rs}(412)$ results from ACA-SIM are more reasonable, even under such strong dust effects.

Figure 15 compares the R_{rs} results at 547 and 667 nm from the 3 algorithms. The consistency among the 3 algorithms is higher in the green and red bands than in the blue bands, although $R_{rs}(547)$ from the NASA Standard is notably lower than that from OC-SMART and ACA-SIM. The 2 NN-based AC algorithms demonstrate the highest agreement ($R^2 = 0.98$ and 0.94 for 547 and 667 nm, respectively), indicating a highly consistent spatial distribution, particularly in turbid coastal waters. In addition, the eddy patterns in the lower left appear much clearer in ACA-SIM $R_{rs}(667)$ than those showing in $R_{rs}(667)$ from the NASA Standard and OC-SMART, and these

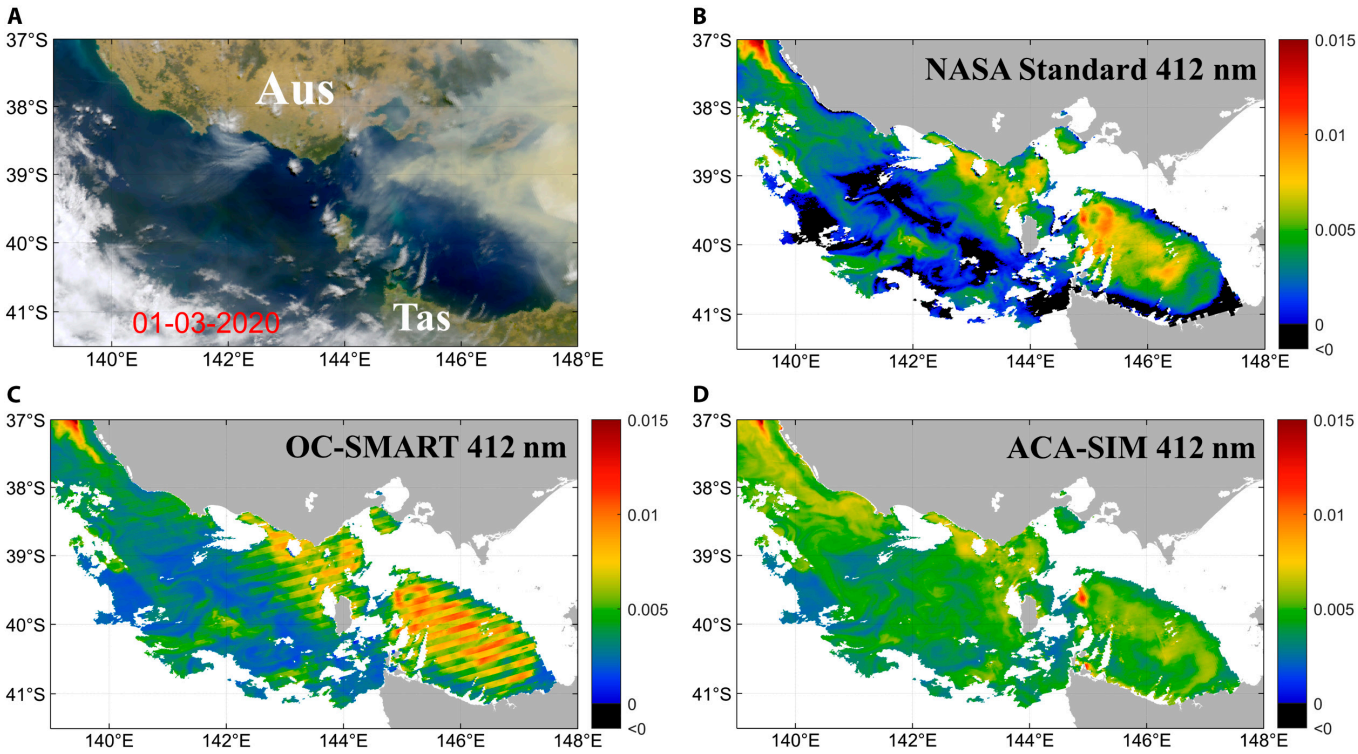

Fig. 16. Results of $R_{rs}(412)$ by the NASA Standard, OC-SMART, and ACA-SIM over the southern coast of Australia collected by MODIS-Aqua on 2020 January 3. (A) RGB true color image. $R_{rs}(412)$ obtained from the NASA Standard (B), OC-SMART (C), and ACA-SIM (D).

eddy patterns are also showing in the $R_{rs}(547)$ images. In addition, both the NASA Standard and OC-SMART were affected by striping artifacts to varying degrees, whereas ACA-SIM showed the potential to reduce the impact of striping artifacts. It is important to note that the striping artifacts are unrelated to the AC algorithms themselves, but rather stem from the quality of the input data. ACA-SIM benefits from the stable true values in its training dataset, which provides it with potential correction capability when dealing with input data affected by striping artifacts.

Bushfires in Australia

The smoke aerosol generated by combustion is one of the most strongly absorbing types of aerosol. From September 2019 to March 2020, a severe bushfire lasted for 6 months in northern and eastern Australia [79], which produced widespread smoke across most of the eastern coastal waters and was transported southeastward by prevailing winds, crossing the Pacific Ocean and reaching South America and the Southern Ocean. Figure 16A shows an RGB composite image from MODIS-Aqua on 2020 January 3 of the Bass Strait between southeastern Australia and Tasmania. The Bass Strait lies to the west of the fire center, where the high concentration of yellow–brown smoke aerosol and its plume are moving eastward, forming a visible thin layer over the strait. Figure 16B to D show the $R_{rs}(412)$ produced by the NASA Standard, OC-SMART, and ACA-SIM, respectively. Again, the NASA Standard obtained much lower $R_{rs}(412)$ (dark blue) for many areas, along with many regions where the values are negative (black), even in the center of the Strait. This further highlights the limitation of the NASA Standard AC algorithm for observations under strong absorbing aerosols.

OC-SMART did not produce negative $R_{rs}(412)$, and the spatial distribution of the resulting $R_{rs}(412)$ appears reasonable, but, again, there are noticeable striping artifacts (see Fig. 16C). This is mainly because the collected data are on the edge of the image, which further amplifies the striping artifact [80,81]. While OC-SMART effectively avoided negative $R_{rs}(412)$ values, it is interesting to see a strong contrast of much higher $R_{rs}(412)$ in the eastern side of the Strait [$R_{rs}(412) \sim 0.01\ sr^{-1}$] than that in the western side of the Strait [$R_{rs}(412) \sim 0.002\ sr^{-1}$], while

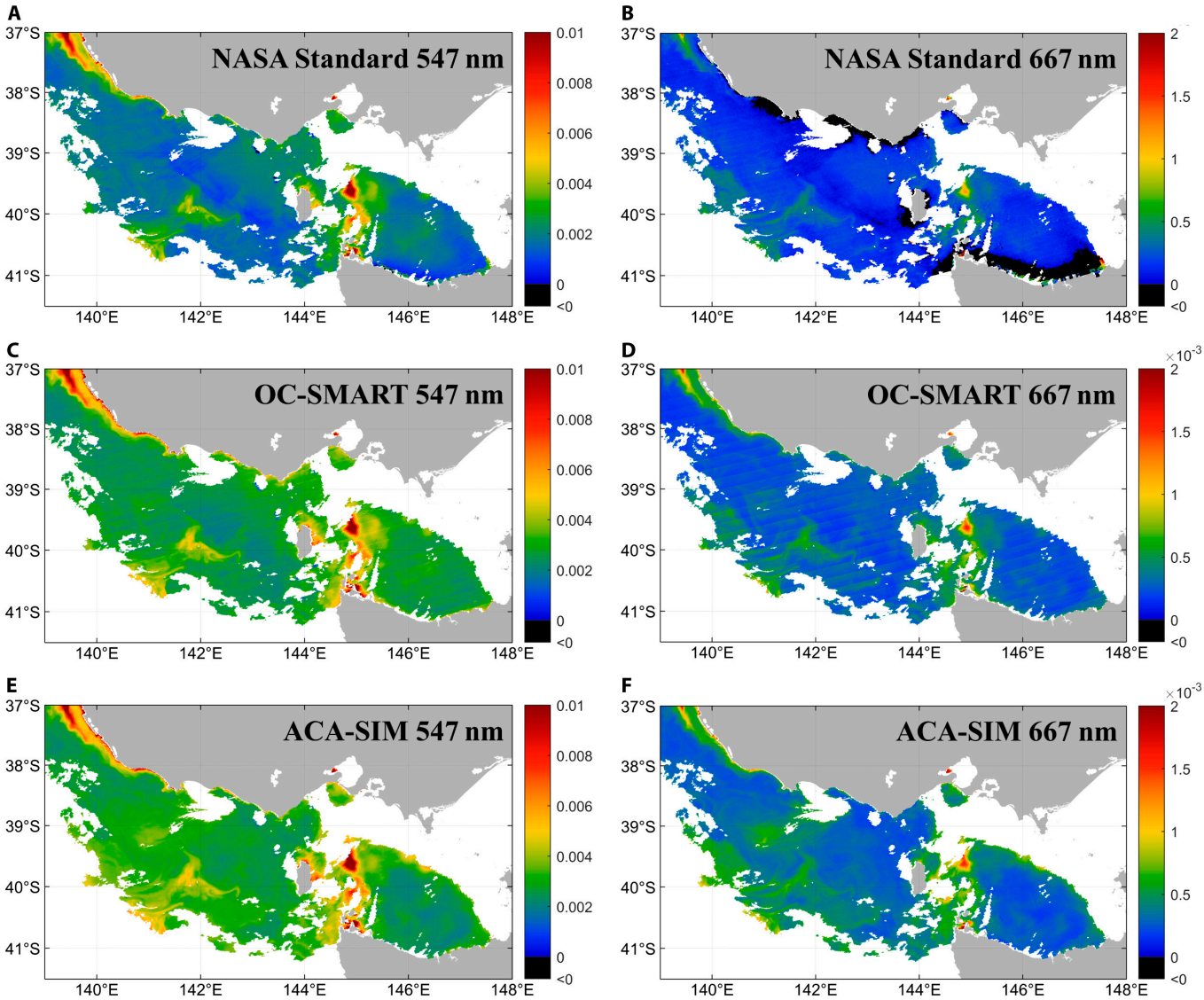

Fig. 17. Results of $R_{rs}(547)$ and $R_{rs}(667)$ by the NASA Standard, OC-SMART, and ACA-SIM for the same scene and date as Fig. 16. (A, C, and E) $R_{rs}(547)$ obtained from the NASA Standard, OC-SMART, and ACA-SIM, respectively; (B, D, and F) $R_{rs}(667)$ obtained from the NASA Standard, OC-SMART, and ACA-SIM, respectively.

possible, but might not be consistent with the dynamic system of such a Strait, especially since there is no such east–west water-mass contrast in the images of R_{rs}(547) and R_{rs}(667) (see Fig. 17C to F), obtained by both OC-SMART and ACA-SIM. In contrast to the results of the NASA Standard and OC-SMART, the R_{rs}(412) from ACA-SIM not only avoided negative values, but there were no visible striping effects, and the spatial distribution is more smooth and uniform, suggesting more reasonable products from ACA-SIM.

Also, as examples for the green and red bands, Fig. 17 shows the R_{rs} values at 547 and 667 nm from the 3 ACs. Similar to that of West Africa (Fig. 14), the NASA Standard produced the lowest R_{rs}(547) among the 3 AC algorithms and negative R_{rs}(667) for nearshore regions (see Fig. 17B). OC-SMART and ACA-SIM, however, produced quite consistent R_{rs} values at these 2 bands, indicating an effectiveness of the data-based AC algorithms. However, ACA-SIM, which used actual MODIS-Aqua measurements to train an NN, clearly showed an advantage in handling sensor-related issues, such as the striping effects.

Discussion

Absorbing aerosol and turbid water

The above compares the performance of the NASA Standard, OC-SMART, and ACA-SIM for measurements under various aerosol loading conditions, including dust, wildfire, and urban/industrial aerosol types (see Figs. 12 to 17). Results indicate that ACA-SIM outperforms the NASA Standard and provides additional improvements over OC-SMART across different scenarios. These results highlight the difficulties of the NASA Standard and OC-SMART in handling these challenging conditions commonly occurring in the coastal environment, whereas a lack of consideration of the vertical distribution of aerosols [82,83] and adjacency effects [34] could cause uncertainties in the AC process. ACA-SIM, on the other hand, as it used actual real-world data in the training process, handles these situations implicitly through the NN's internal learning processes and achieves better results.

The dynamic range and representativeness of the training dataset always play a crucial role in any DD algorithm. Therefore, we analyzed the aerosol characteristics encountered by the training data used in this study. The aerosol classification algorithm [84] developed based on AERONET measurements was employed for this analysis. This algorithm classifies aerosols into biomass burning (or smoke), dust, mixed absorbing aerosols, marine, continental, and mixed nonabsorbing aerosols based on the Angstrom exponent (AE) and single scattering albedo (SSA). Simply put, AE and SSA, respectively, characterize the particle size distribution and radiative absorption properties of aerosols [85]. When combined, they can effectively differentiate between distinct aerosol types [85–87]. The AE and SSA data were sourced from the AERONET Version 3 Inversion Products [88,89], with Fig. 18A showing the distribution of aerosol types within the ACA-SIM training set. Among them, the nonabsorbing aerosols—marine, continental, and mixed aerosols—account for 3.5%, 14.1%, and 19.1%, respectively. In contrast, the absorbing aerosols, including smoke, dust, and mixed absorbing aerosols, make up 16.4%, 7.8%, and 39.1%, respectively. Overall, ~63% of the data were obtained under absorbing aerosol conditions, which provides a strong reason for the excellent performance of ACA-SIM for measurements under such conditions.

Additionally, there is a strong presence of dust aerosols [90] in our training dataset, with SSA values reaching as low as 0.34, indicating extremely strong absorbing characteristics. Furthermore, in coastal areas, mixtures of multiple aerosol types are very common than the scenario of a single aerosol type [91]. In the ACA-SIM training dataset, nearly 40% of the data were classified as mixed absorbing aerosols, indicating that ACA-SIM is suitable for regions with complex atmospheric conditions where aerosols of multiple types often coexist.

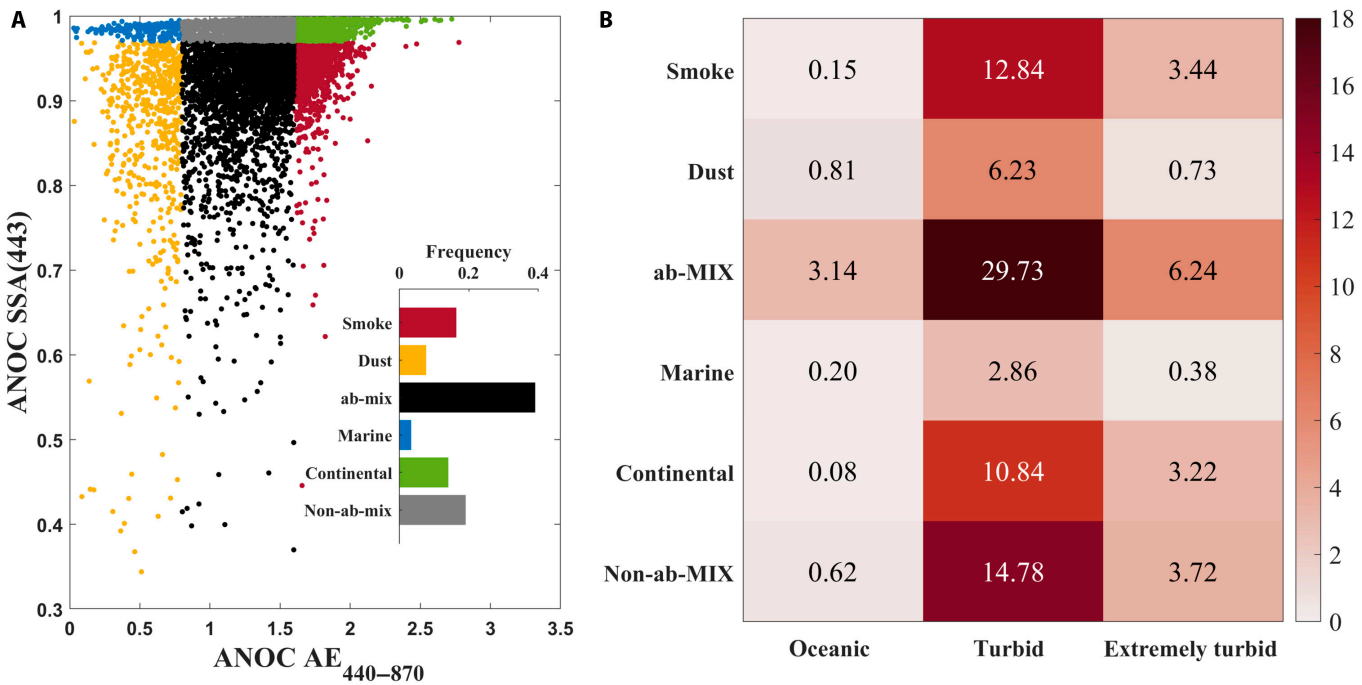

Fig. 18. Dynamic ranges of atmosphere and water properties of the ACA-SIM training set. (A) Aerosol classification and frequency distribution. (B) Heatmap of coupled water types and aerosol types.

In addition, a common scenario in coastal regions is the coexistence of turbid waters and absorbing aerosols. By applying a quality assurance (QA) system [92] to classify water types into oceanic (QA cluster = 1 to 4), turbid (QA cluster = 5 to 18), and extremely turbid waters (QA cluster = 19 to 23), a coupling matrix of water and aerosol types was generated (see Fig. 18B). It is found that the ACA-SIM training set includes ~30% of turbid water mixed with absorbing aerosols, a prevalent condition in coastal environments. On the other hand, more than 5% and 17% of the data were collected under clear oceanic and extremely turbid water conditions, respectively. These distributions indicate a broad dynamic range and the representativeness of the training data, supporting the applicability of ACA-SIM in complex coastal environments.

To further explore the potential of ACA-SIM in highly turbid waters, we applied the 3 AC algorithms to a MODIS-Aqua image from 2019 November 11 (see Fig. 19). This image covers the Subei Bank in the southwestern Yellow Sea and part of the Yangtze River Delta. Due to the strong sediment discharge from the Yangtze River estuary and strong resuspension of the shallow Subei Bank, these regions are characterized by extremely turbid waters (see Fig. 19A).

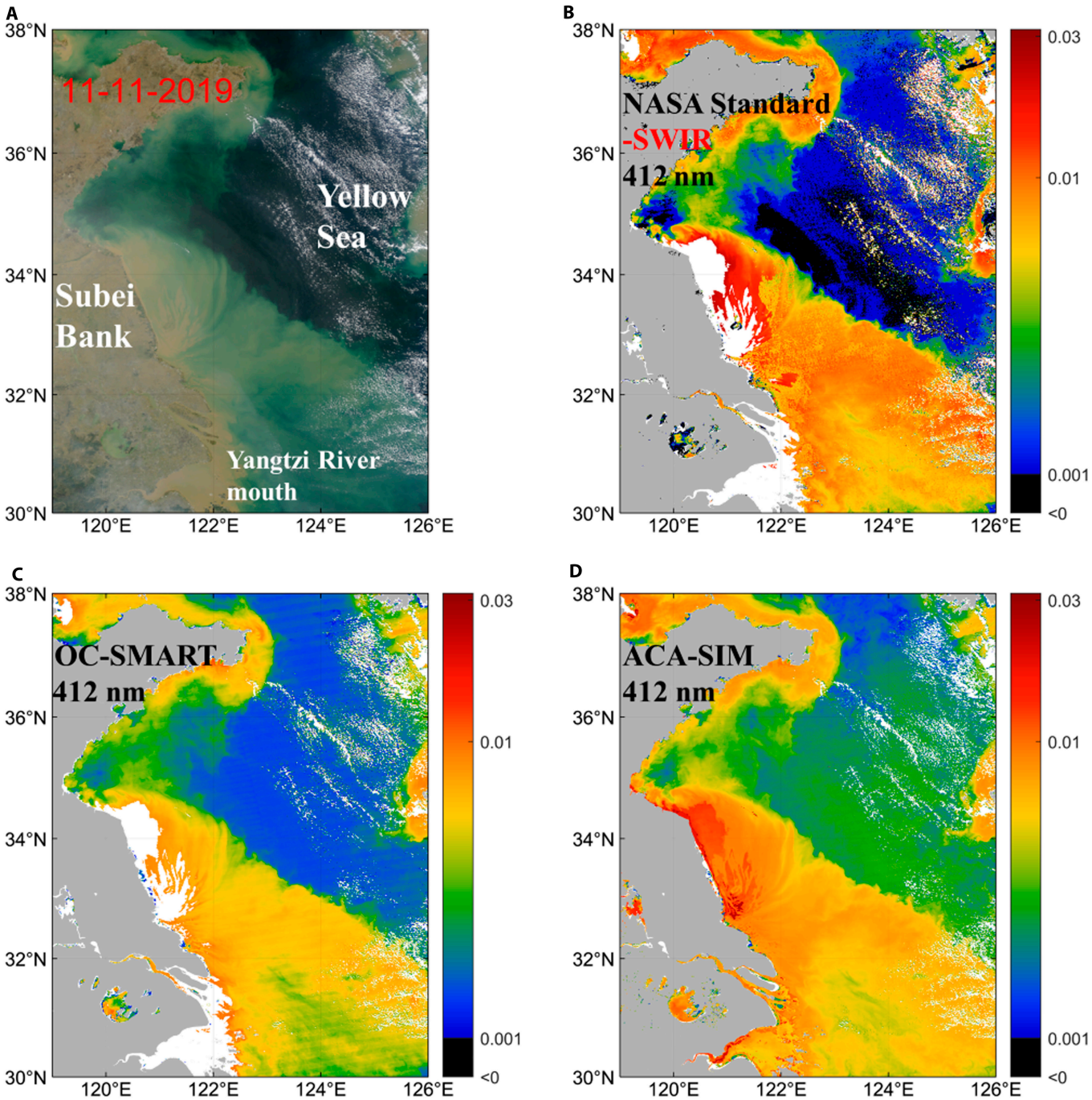

Fig. 19. Results of R_{rs}(412) by NASA Standard-SWIR (B), OC-SMART (C), and ACA-SIM (D) over the Subei Bank collected by MODIS-Aqua on 2019 November 11. (A) RGB true color image. R_{rs}(412) obtained from NASA Standard-SWIR (B), OC-SMART (C), and ACA-SIM (D), using a logarithmic scale.

To process this MODIS-Aqua measurement, the NASA Standard utilized the SWIR scheme [18], termed NASA Standard-SWIR here, which enables more valid retrievals under such conditions compared to the iterative NIR scheme. For this exceptional water body, all 3 AC algorithms showed a similar spatial distribution of R_{rs}(412) (see Fig. 19B to D), but NASA Standard-SWIR produced some negative R_{rs}(412) in the central ocean region. For the extremely turbid Subei Bank, R_{rs}(412) values by NASA Standard-SWIR and ACA-SIM appear quite close, but the results from OC-SMART are somewhat lower. These results suggest that ACA-SIM is applicable to such extremely turbid waters, although the longest wavelength involved is 1,240 nm, rather the SWIR bands required by NASA Standard-SWIR. It is necessary to point out that, unlike the comparisons presented in the "Results and Evaluations" section, we here limited the mask of cloud only for results from ACA-SIM, in order to explore the upper limit of ACA-SIM in highly turbid waters.

Striping artifacts

The primary source of striping artifacts lies in calibration inconsistencies among multiple detectors within different bands of the sensor, the reflectance differences of the scan mirror's double-sided surfaces, and the slightly geometric angle variations between the different detectors during imaging [53]. This phenomenon is particularly common in sensors with "Whisk-broom scanner" imaging methods, such as MODIS and VIIRS. Moreover, striping effects may become more pronounced at larger scanning angles [80], corresponding to the image edges. In the ocean color bands of MODIS-Aqua, striping artifacts caused by the scan mirror are the most common, while those resulting from geometric angle variations between the different detectors are also relatively common under extreme conditions.

Striping strongly impacts the accuracy of AC algorithms and can propagate errors into downstream ocean color products [81]. In our study, both the NASA Standard and OC-SMART exhibited varying degrees of striping in image products. It is important to note that this is not a flaw in the NASA Standard or OC-SMART algorithms themselves but rather a consequence of the input data quality.

Unlike these 2 approaches, as ACA-SIM used striping-affected MODIS-Aqua data in the training process, it inherently possesses the capabilities to minimize the striping effects through learning. Specifically, because of the use of 2 mirrors (both sides) of the rotating scan mirror to obtain images, all MODIS-Aqua data inherently contain real scan mirror-caused striping artifacts. The only difference lies in the severity of the striping. To investigate the distribution of more severe scan mirror-caused striping artifacts in the ACA-SIM training set, a striping detection algorithm [93] was employed. Severe striping was identified using a threshold where the relative deviation of brightness exceeded ±0.008. Details of this algorithm are described in Xiang and Liu [93]. Figure 20 shows the proportion of scan mirror-caused striping at each station within the training set. It is found that 21.5% of the data were acquired under the influence of severe scan mirror-caused striping. This explains why ACA-SIM can effectively minimize the striping artifacts. Also, note that similar striping effects are observed at all stations, indicating that ACA-SIM has learned a wide range of water properties under striping effects, thereby enhancing ACA-SIM's capability to process measurements from various coastal waters. While we have demonstrated the potential of ACA-SIM to mitigate striping artifacts, it is worth noting that it does not yet completely eliminate them. In the ACA-SIM results presented in this study, slight traces of striping are still visible. Therefore, integrating more data samples and incorporating effective spatial information into the model remain necessary.

Conclusion

In satellite ocean color remote sensing, the overwhelming contribution from the atmosphere has made it a challenge to remove atmospheric effects, thus obtaining accurate water-leaving signals, especially for coastal waters due to complex combinations of water and aerosol properties. In this study, aided by greatly advanced NNs, we demonstrated that using the matchup data between satellite TOA measurements and "ground truth" R_{rs} can substantially improve the quality of R_{rs} of various coastal environments. Compared to the NASA Standard, the developed AC algorithm in this effort, ACA-SIM, avoided the generation of negative R_{rs} values, improved accuracy across the blue to red bands, and handled complex coastal water conditions much better. Its ability to remove striping artifacts further enhances the reliability of the derived R_{rs} products.

This advancement and effectiveness in AC stem from the fact that satellite data represent what happens in the real world, which not only covers various water and aerosol properties, as well as measurement geometries and land adjacent effects, but also includes the characteristics of the satellite sensor making the measurements. These combinations make it extremely difficult to accurately simulate the radiance measured by a satellite sensor via coupled atmosphere–ocean models. While we observe

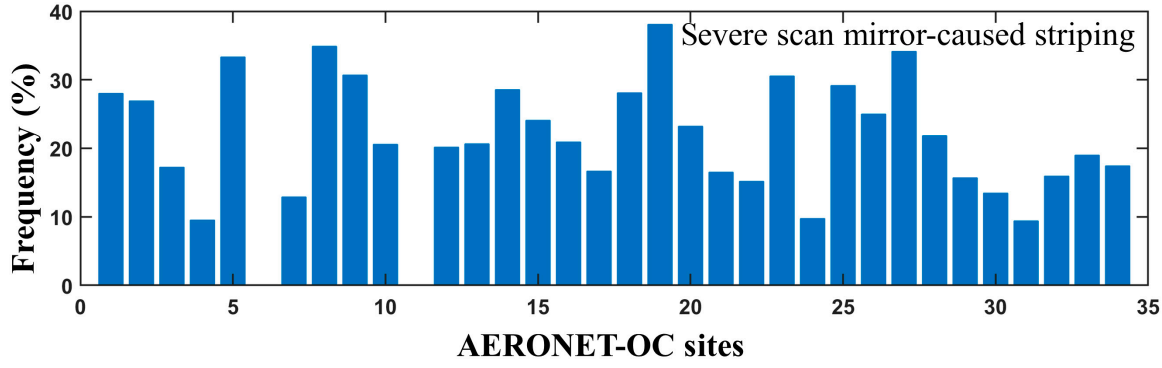

Fig. 20. Proportion of severe striping at each station in the training set.

the excellent performance of the present version of ACA-SIM, as is always the case with algorithm development, we expect ACA-SIM to improve further after more AERONET-OCs are installed in regions beyond the existing water–atmosphere combinations, thereby expanding the training data boundary. In the meantime, it is necessary to extend ACA-SIM to other satellite sensors and, therefore, to enhance the coverage and consistency of coastal aquatic environments by multiple satellites. However, it is necessary to point out that this study focused on optically deep waters in coastal regions. For the AC of optically shallow waters, ACA-SIM remains conceptually applicable, but it will require an extremely large and representative training set due to the wide range of bottom substrates and depths.

Acknowledgments

We thank NASA for distributing MODIS ocean color products, the many AERONET-OC sites, and the SeaBASS program for providing valuable measurements. We also appreciate the support of the NSFC Shiptime Sharing Project (data and samples collected onboard R/V "Haike 68" and "Shiyan 3" during cruise NORC2023-07 supported by NSFC Project #42249907) and Xiamen University for facilitating the collection of field measurements. Constructive comments and suggestions from Dr. Giuseppe Zibordi and an anonymous reviewer are greatly appreciated.

Funding: We are grateful for the financial support from the National Key Research and Development Program of China (2022YFC3104901), the National Natural Science Foundation of China (#42430107), Fujian Ocean Innovation Laboratory, Fujian Satellite Data Development, Co., Ltd., and Fujian Haisi Digital Technology Co., Ltd.

Author contributions: Z.L. conceptualized the study and revised and finalized the manuscript. X.L. conducted the study, generated the dataset, developed the neural networks, analyzed the data, and drafted the manuscript. J.X. helped with MODIS-Aqua data preparation and neural-network improvement. T.W. helped with data preparation and result analysis. L.Z. helped with data preparation and result analysis. D.W. helped with data preparation and result analysis. X.Y. helped with the study and revised the manuscript. S.S. helped with the study and revised the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data Availability

Data used in this study are available for download from the NASA Ocean Color site (https://oceandata.sci.gsfc.nasa.gov/), the AERONET site (https://aeronet.gsfc.nasa.gov/), and the SeaBASS site (https://seabass.gsfc.nasa.gov/).

Appendix A: Shift AERONET-OC bands to MODIS bands

As AERONET-OC sites have different band settings compared to that of MODIS, the spectral bands of ANOC-R_{rs} have to be shifted to match that of MODIS in order to develop an NN-based algorithm to process MODIS data. There could be many ways for this band shift [94–96]; in this study, similar to Ref. [94], we used bio-optical modeling to meet the goal. Specifically, we first constructed a hyperspectral R_{rs} from the multi-band ANOC-R_{rs}, then through spectral convolution to get the R_{rs} of MODIS bands. For the construction of hyperspectral R_{rs}, it

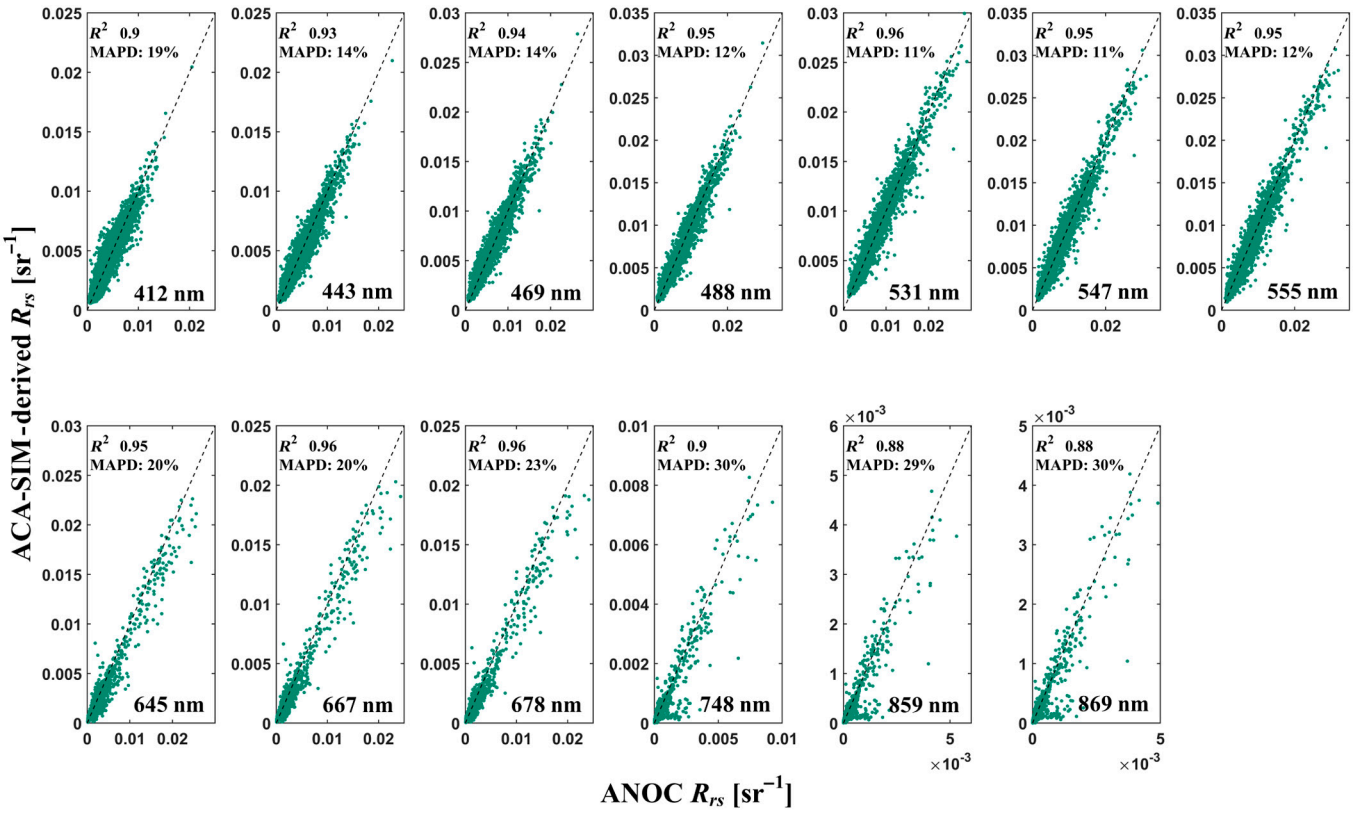

Fig. A1. Comparison between R_{rs} derived from ACA-SIM and ANOC-R_{rs} when the training is completed (n = 7,054).

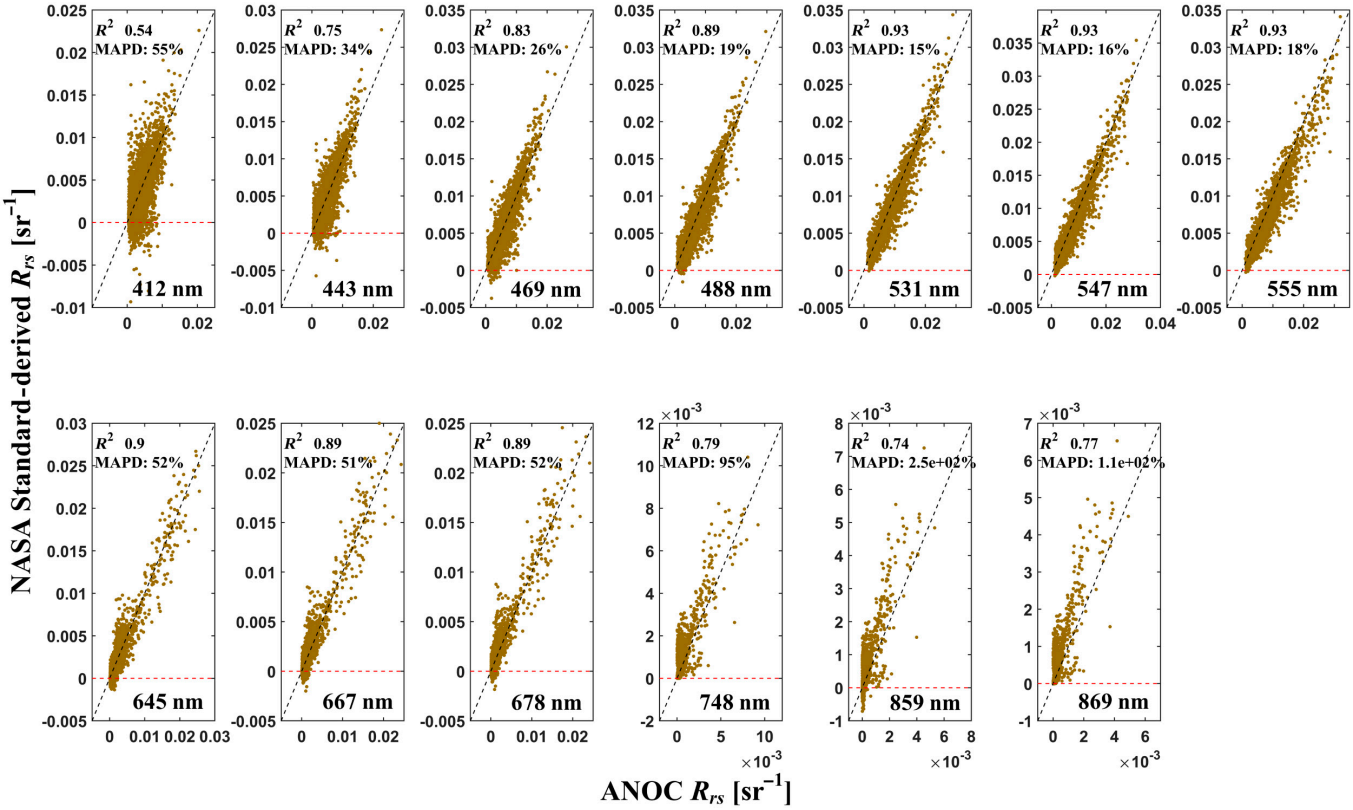

Fig. A2. Comparison between R_{rs} derived from the NASA Standard and ANOC-R_{rs} of the training data (n = 7,054).

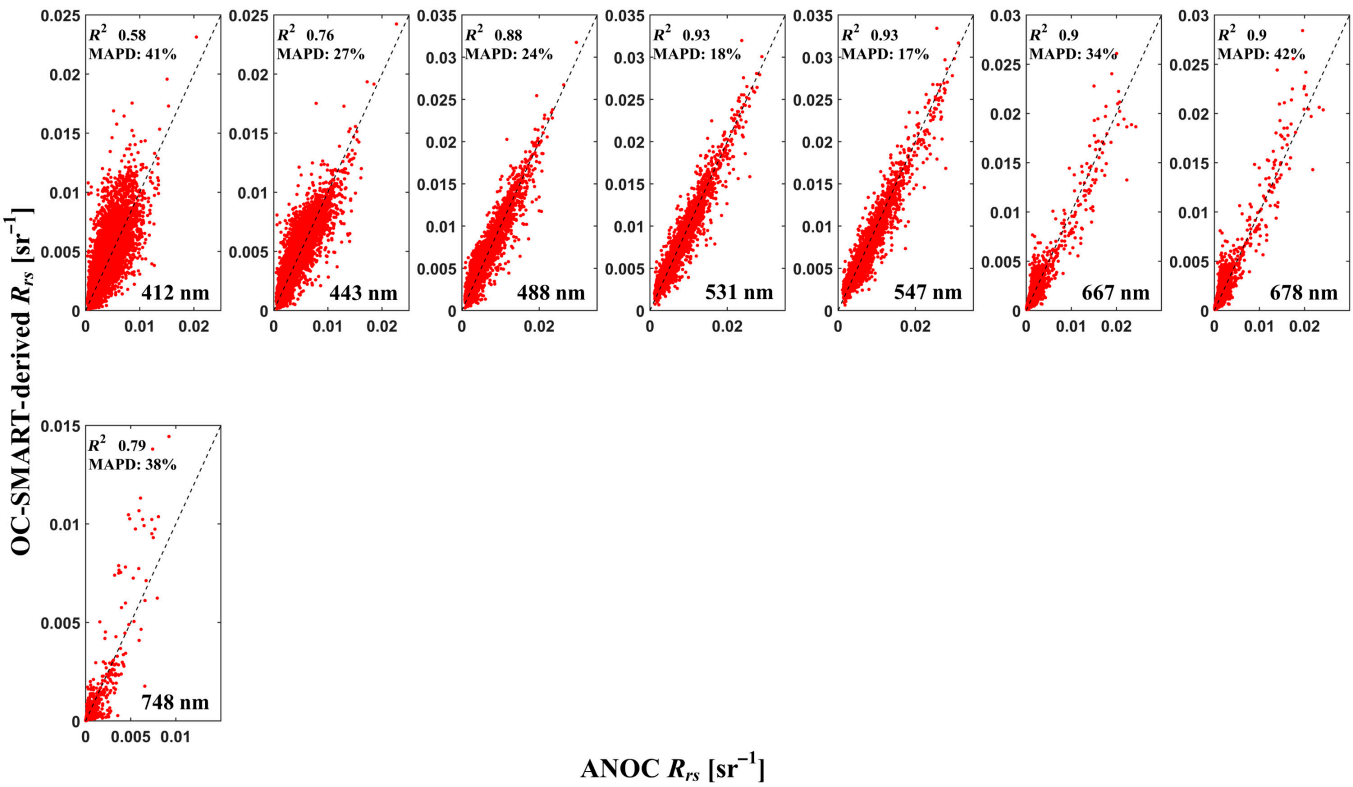

Fig. A3. Comparison between R_{rs} derived from OC-SMART and ANOC-R_{rs} of the training data (n = 7,054).

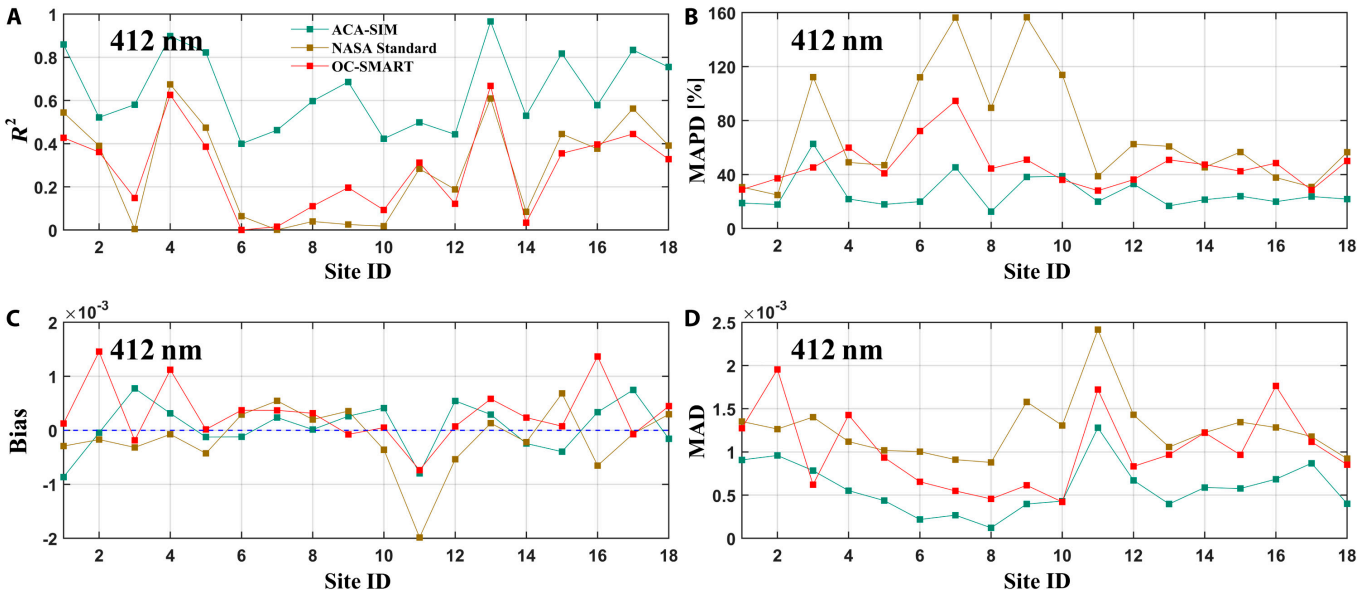

Fig. A4. Statistical results of the leave-one-site-out validation at 412 nm. (A) Range of R^2, (B) Range of MAPD, (C) Range of Bias, (D) Range of MAD.

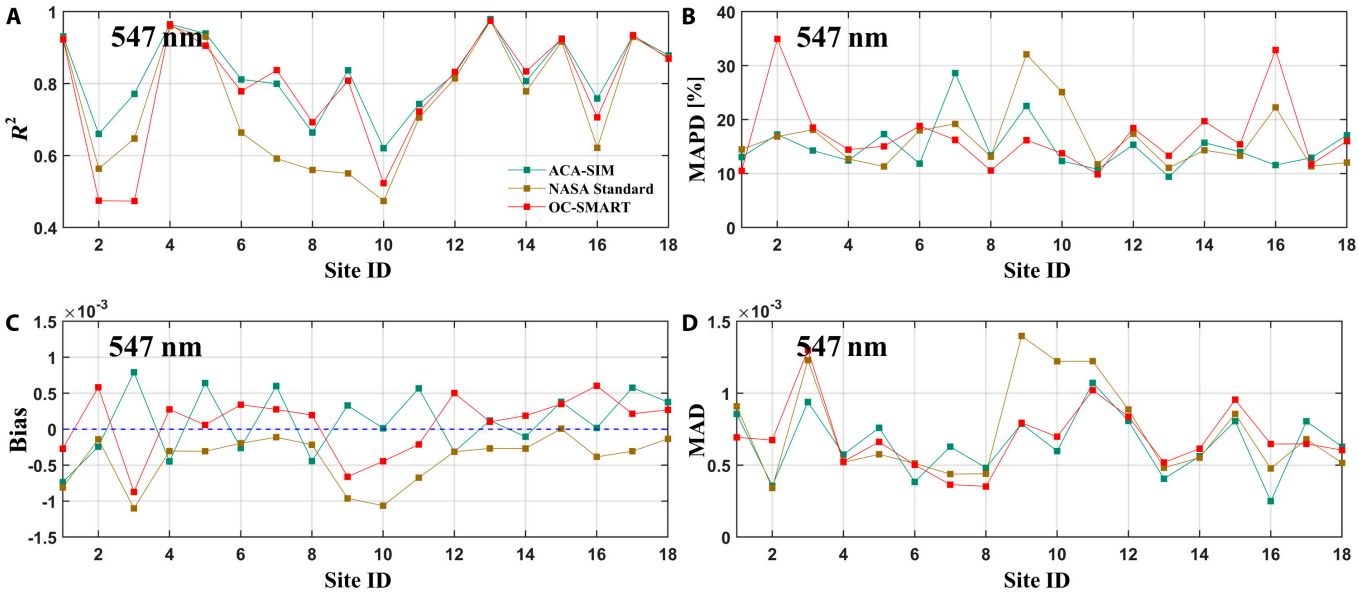

Fig. A5. Statistical results of the leave-one-site-out validation at 547nm. (A) Range of R^2, (B) Range of MAPD, (C) Range of Bias, (D) Range of MAD.

was achieved from bio-optical models developed and widely utilized in the past decades. Generally, for optically deep waters, R_{rs} can be summarized as [97,98]

$$R_{rs}(\lambda) = f\left(a(\lambda), b_b(\lambda)\right). \tag{A1}$$

Here, a and b_b are the total absorption and backscattering coefficients of a water body, which can be further expressed as for turbid coastal waters [99],

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_g(\lambda) + a_{sed}(\lambda), \tag{A2}$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda), \tag{A3}$$

where a_w and b_{bw} are the absorption and backscattering coefficients of seawater and pure seawater. a_{ph}, a_g, a_{sed}, and b_{bp} represent the absorption coefficient of phytoplankton, gelbstoff, suspended sediments, and the backscattering coefficient of particles, respectively. While a_w and b_{bw} are considered constants [100,101], the other components are modeled as

$$a_{ph}(\lambda) = a_{ph}(440)a_{ph}^{+}(\lambda), \tag{A4}$$

$$a_g(\lambda) = a_g(440)e^{-0.014(\lambda-440)}, \tag{A5}$$

$$a_{sed}(\lambda) = A_{sed}\, a_{sed}^{+}(\lambda) + B_{sed}, \tag{A6}$$

$$b_{bp}(\lambda) = b_{bp}(440)\left(\frac{440}{\lambda}\right)^{Y}. \tag{A7}$$

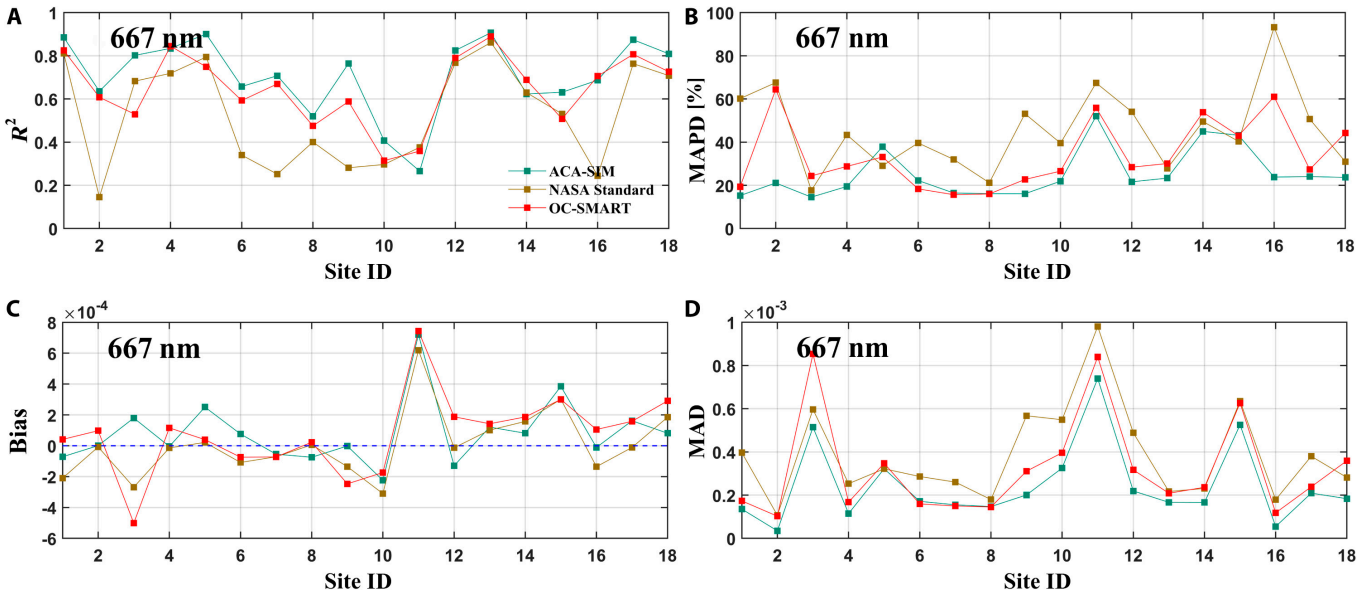

Fig. A6. Statistical results of the leave-one-site-out validation at 667 nm. (A) Range of R^2, (B) Range of MAPD, (C) Range of Bias, (D) Range of MAD.

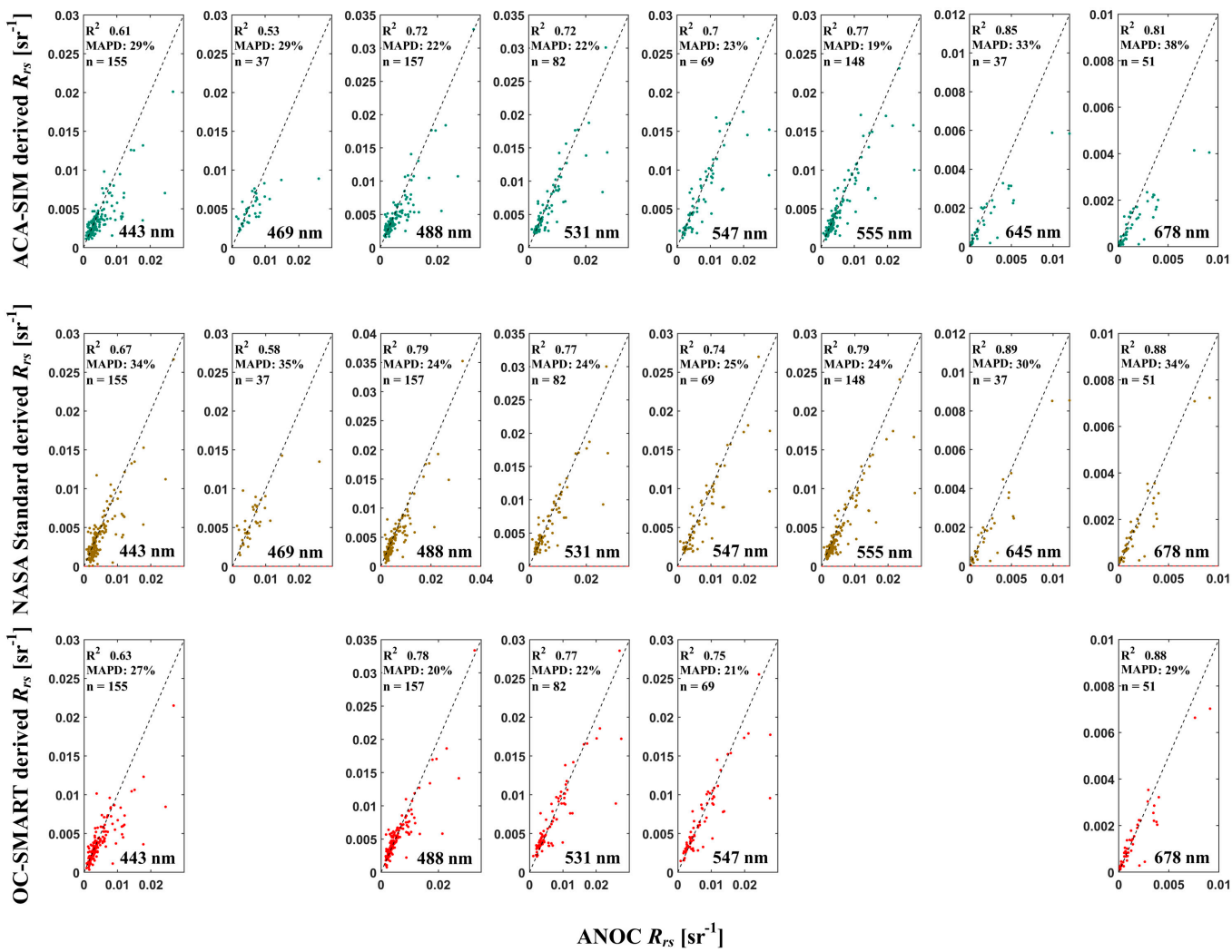

Fig. A7. Scatterplot of retrieved R_{rs} by ACA-SIM (green), the NASA Standard (brown), and OC-SMART (red) algorithms compared with in situ measurements.

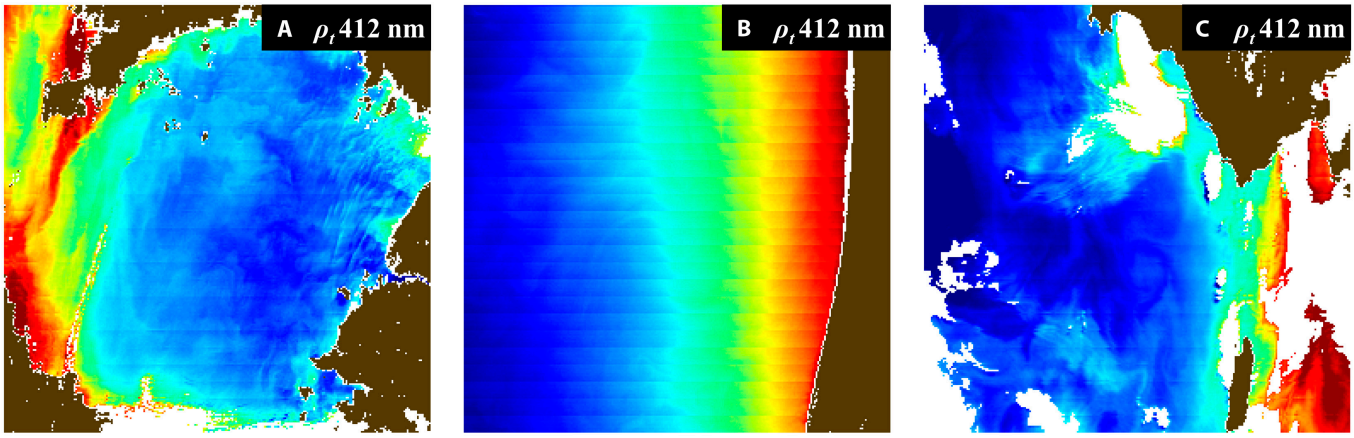

Fig. A8. Unreprojected single-band raw image of ρ_t at 412 nm. (A) to (C) show subregions of the images corresponding to the sections: (A) 5.4.1 Bohai and Yellow Sea; (B) 5.4.2 West Coast of North Africa; and (C) 5.4.3 Bushfires in Australia.

where $a^+_{ph}(\lambda)$ and $a^+_{sed}(\lambda)$ are the spectral shape of phytoplankton and sediment absorption coefficient, respectively, with $a^+_{ph}(\lambda)$ provided in IOCCG Report #5 [8] and $a^+_{sed}(\lambda)$ provided in Ref. [99]). Separatelwy, as in Lee et al. [102], a term (Δ) to account for the residual error in the ANOC-R_{rs} was included. Thus, there are 7 unknowns [$a_{ph}(440)$, $a_g(440)$, A_{sed}, B_{sed}, $b_{bp}(440)$, Y, and Δ] for an R_{rs} spectrum, which can be estimated through spectral optimization [56,103].

After the 7 variables are estimated, apply the values of the 7 variables back to Eqs. A1 to Eqs. A7, and hyperspectral R_{rs} spectrum (the R_{rs} in Eq. 8 with a wave sign on head) can then be generated. Finally, the reconstructed hyperspectral is convolved to the MODIS-Aqua bands (or the bands of any specified satellite sensor) using the sensor's spectral response function (SRF),

$$R_{rs}(\lambda) = \int_{\lambda_1}^{\lambda_2} \widetilde{R}_{rs}(\lambda)\,\mathrm{SRF}(\lambda)d\lambda \qquad \text{(A8)}$$

As examples, Fig. 2 shows re-constructed hyperspectral R_{rs} from the multi-band R_{rs}, and then the R_{rs} at the MODIS bands. Analysis using hyperspectral data indicates that the average error in the band-shifted R_{rs} is ~6%, a quality sufficient to train ACA-SIM.

Appendix B: Performance of the 3 AC algorithms with the training dataset

Scatterplots (Figs. A1 to A3) between the AC-algorithm-derived R_{rs} and ANOC-R_{rs} of the training data are presented here, which include ACA-SIM, the NASA Standard, and OC-SMART. Since this dataset was used to train ACA-SIM, it is not an independent evaluation of this algorithm, just a reference. For both the NASA Standard and OC-SMART, it is an evaluation based on independent data from coastal waters.

Appendix C: Results of representative blue–green–red bands (412, 547, and 667 nm) in leave-one-site-out validation

To supplement the leave-one-site-out evaluation, here (Figs. A4 to A6) the breakdown results of the 3 algorithms at the blue (412 nm), green (547 nm), and red (667 nm) are presented. It is found that, regardless of which site was left out for validation, ACA-SIM shows better overall statistical measures, especially for the blue band.

Appendix D: Performance on the ship-based measurements

To supplement the comparisons shown in Fig. 11, scatterplots (Fig. A7) between algorithm-derived R_{rs} and in situ measurements of coastal waters for the other bands in the blue–red domain are presented here.

Appendix E: Single-band raw image of ρ_t at 412 nm

To supplement the comparison shown in Figs. 12 to 16, the unprojected single-band raw images of ρ_t at 412 nm are presented here (Fig. A8). It can be seen that all 3 raw images are affected by horizontal striping artifacts to varying degrees.

References

1. Dai Y, Yang S, Zhao D, Hu C, Xu W, Anderson DM, Li Y, Song X-P, Boyce DG, Gibson L, et al. Coastal phytoplankton blooms expand and intensify in the 21st century. *Nature*. 2023;615(7951):280–284.
2. Breitburg D, Levin LA, Oschlies A, Grégoire M, Chavez FP, Conley DJ, Garçon V, Gilbert D, Gutiérrez D, Isensee K, et al. Declining oxygen in the global ocean and coastal waters. *Science*. 2018;359(6371):Article eaam7240.
3. Cloern JE, Abreu PC, Carstensen J, Chauvaud L, Elmgren R, Grall J, Greening H, Johansson JO, Kahru M, Sherwood ET, et al. Human activities and climate variability drive fast-paced change across the world's estuarine-coastal ecosystems. *Glob Chang Biol*. 2016;22(2):513–529.
4. Mouw CB, Greb S, Aurin D, DiGiacomo PM, Lee Z, Twardowski M, Binding C, Hu C, Ma R, Moore T, et al. Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sens Environ*. 2015;160:15–30.
5. IOCCG. Atmospheric correction for remotely-sensed ocean colour products. In: Wang M, editor. *Reports of International Ocean-Colour Coordinating Group*. Dartmouth (Canada): IOCCG; 2010. p. 83.

6. Gordon HR, Wang M. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. *Appl Opt*. 1994;33(3):443–452.
7. IOCCG. Remote sensing of ocean colour in coastal, and other optically-complex, waters. In: Sathyendranath S, editor. *Reports of International Ocean-Colour Coordinating Group*. Dartmouth (Canada): IOCCG; 2000. p. 140.
8. IOCCG. Remote sensing of inherent optical properties: Fundamentals, tests of algorithms, and applications. In: Lee Z, editor. *Reports of International Ocean-Colour Coordinating Group*. Dartmouth (Canada): IOCCG; 2006. p. 122.
9. Gordon HR. Atmospheric correction of ocean color imagery in the earth observing system era. *J Geophys Res Atmos*. 1997;102(D14):17081–17106.
10. Feng H, Vandemark D, Campbell JW, Holben BN. Evaluation of MODIS Ocean colour products at a Northeast United States coast site near the Martha's Vineyard coastal observatory. *Int J Remote Sens*. 2008;29(15):4479–4497.
11. Hlaing S, Harmel T, Gilerson A, Foster R, Weidemann A, Arnone R, Wang M, Ahmed S. Evaluation of the VIIRS Ocean color monitoring performance in coastal regions. *Remote Sens Environ*. 2013;139:398–414.
12. Dogliotti AI, Ruddick K, Guerrero R. Seasonal and inter-annual turbidity variability in the Río de la Plata from 15 years of MODIS: El Niño dilution effect. *Estuar Coast Shelf Sci*. 2016;182:27–39.
13. Wei J, Yu X, Lee Z, Wang M, Jiang L. Improving low-quality satellite remote sensing reflectance at blue bands over coastal and inland waters. *Remote Sens Environ*. 2020;250: Article 112029.
14. Siegel DA, Wang M, Maritorena S, Robinson W. Atmospheric correction of satellite ocean color imagery: The black pixel assumption. *Appl Opt*. 2000;39(21):3582–3591.
15. Wang J, Wang Y, Lee Z, Wang D, Chen S, Lai W. A revision of NASA SeaDAS atmospheric correction algorithm over turbid waters with artificial neural networks estimated remote-sensing reflectance in the near-infrared. *ISPRS J Photogramm Remote Sens*. 2022;194:235–249.
16. Ruddick KG, Ovidio F, Rijkeboer M. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl Opt*. 2000;39(6):897–912.
17. Bailey SW, Franz BA, Werdell PJ. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Opt Express*. 2010;18(7):7521–7527.
18. Wang M, Shi W. The NIR-SWIR combined atmospheric correction approach for MODIS Ocean color data processing. *Opt Express*. 2007;15(24):15722–15733.
19. Gao BC, Montes MJ, Li RR, Dierssen HM, Davis CO. An atmospheric correction algorithm for remote sensing of bright coastal waters using MODIS land and ocean channels in the solar spectral region. *IEEE Trans Geosci Remote Sens*. 2007;45(6):1835–1843.
20. He Q, Chen C. A new approach for atmospheric correction of MODIS imagery in turbid coastal waters: A case study for the Pearl River estuary. *Remote Sens Lett*. 2014;5(3):249–257.
21. He X, Bai Y, Pan D, Tang J, Wang D. Atmospheric correction of satellite ocean color imagery using the ultraviolet wavelength for highly turbid waters. *Opt Express*. 2012;20(18):20754–20770.
22. Vanhellemont Q, Ruddick K. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. *Remote Sens Environ*. 2018;216:586–597.
23. Vanhellemont Q. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sens Environ*. 2019;225:175–192.
24. Brajard J, Moulin C, Thiria S. Atmospheric correction of SeaWiFS Ocean color imagery in the presence of absorbing aerosols off the Indian coast using a neuro-variational method. *Geophys Res Lett*. 2008;35(20):Article 2008GL035179.
25. Li H, He X, Bai Y, Shanmugam P, Park Y-J, Liu J, Zhu Q, Gong F, Wang D, Huang H. Atmospheric correction of geostationary satellite ocean color data under high solar zenith angles in open oceans. *Remote Sens Environ*. 2020;249:Article 112022.
26. Schiller H, Doerffer R. Neural network for emulation of an inverse model operational derivation of case II water properties from MERIS data. *Int J Remote Sens*. 1999;20(9):1735–1746.
27. Schroeder T, Behnert I, Schaale M, Fischer J, Doerffer R. Atmospheric correction algorithm for MERIS above case-2 waters. *Int J Remote Sens*. 2007;28(7):1469–1486.
28. Schroeder T, Schaale M, Lovell J, Blondeau-Patissier D. An ensemble neural network atmospheric correction for Sentinel-3 OLCI over coastal waters providing inherent model uncertainty estimation and sensor noise propagation. *Remote Sens Environ*. 2022;270:Article 112848.
29. Song Z, He X, Bai Y, Dong X, Wang D, Li T, Zhu Q, Gong F. Atmospheric correction of absorbing aerosols for satellite ocean color remote sensing over coastal waters. *Remote Sens Environ*. 2023;290:Article 113552.
30. Wang M, Fan D, He H, Zeng Y, Fu B, Liang T, Zhang X, Hu W. An atmospheric correction method for Himawari-8 imagery based on a multi-layer stacking algorithm. *Eco Inform*. 2025;85:Article 103001.
31. Lee Z, Wang T, Zhao L, Wang D, Ye X, Shang S, Yu X. Cross-satellite atmospheric correction for consistent remote sensing reflectance from multiple ocean color satellites: Concept and demonstrations. *J Remote Sens*. 2024;4:Article 0302.
32. Fan Y, Li W, Gatebe CK, Jamet C, Zibordi G, Schroeder T, Stamnes K. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sens Environ*. 2017;199:218–240.
33. Fan Y, Li W, Chen N, Ahn J-H, Park Y-J, Kratzer S, Schroeder T, Ishizaka J, Chang R, Stamnes K. OC-SMART: A machine learning based data analysis platform for satellite ocean color sensors. *Remote Sens Environ*. 2021;253: Article 112236.
34. Feng L, Hu C. Land adjacency effects on MODIS aqua top-of-atmosphere radiance in the shortwave infrared: Statistical assessment and correction. *J Geophys Res Oceans*. 2017;122(6):4802–4818.
35. Zibordi G, Holben BN, Talone M, D'Alimonte D, Slutsker I, Giles DM, Sorokin MG. Advances in the ocean color component of the aerosol robotic network (AERONET-OC). *J Atmos Ocean Technol*. 2021;38(4):725–746.
36. Zibordi G, Berthon J-F, Mélin F, D'Alimonte D, Kaitala S. Validation of satellite ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic Proper and Gulf of Finland. *Remote Sens Environ*. 2009;113(12):2574–2591.
37. Jamet C, Loisel H, Kuchinke CP, Ruddick K, Zibordi G, Feng H. Comparison of three SeaWiFS atmospheric

correction algorithms for turbid waters using AERONET-OC measurements. *Remote Sens Environ*. 2011;115(8):1955–1965.
38. Pahlevan N, Mangin A, Balasubramanian SV, Smith B, Alikas K, Arai K, Barbosa C, Bélanger S, Binding C, Bresciani M, et al. ACIX-aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. *Remote Sens Environ*. 2021;258:Article 112366.
39. Mélin F. Validation of ocean color remote sensing reflectance data: Analysis of results at European coastal sites. *Remote Sens Environ*. 2022;280:Article 113153.
40. Thuillier G, Hers M, Simon PC, Labs D, Mandel H, Gillotay D. 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*. 1998;35(4):689.
41. Gordon HR, Brown JW, Evans RH. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone color scanner. *Appl Opt*. 1988;27(5):862–871.
42. Wang M. The Rayleigh lookup tables for the SeaWiFS data processing: Accounting for the effects of ocean surface roughness. *Int J Remote Sens*. 2002;23(13):2693–2702.
43. Wang M. A refinement for the Rayleigh radiance computation with variation of the atmospheric pressure. *Int J Remote Sens*. 2005;26(24):5651–5663.
44. Frouin R, Schwindling M, Deschamps P-Y. Spectral reflectance of sea foam in the visible and near-infrared: In situ measurements and remote sensing implications. *J Geophys Res Oceans*. 1996;101(C6):14361–14371.
45. Moore KD, Voss KJ, Gordon HR. Spectral reflectance of whitecaps: Their contribution to water-leaving radiance. *J Geophys Res Oceans*. 2000;105(C3):6493–6499.
46. Wang M, Bailey SW. Correction of sun glint contamination on the SeaWiFS Ocean and atmosphere products. *Appl Opt*. 2001;40(27):4790–4798.
47. Gao B-C, Montes MJ, Ahmad Z, Davis CO. Atmospheric correction algorithm for hyperspectral remote sensing of ocean color from space. *Appl Opt*. 2000;39(6):887–896.
48. Ahmad Z, Franz BA, McClain CR, Kwiatkowska EJ, Werdell J, Shettle EP, Holben BN. New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. *Appl Opt*. 2010;49(29):5545–5560.
49. Stamnes K, Hamre B, Stamnes S, Chen N, Fan Y, Li W, Lin Z, Stamnes J. Progress in forward-inverse modeling based on radiative transfer tools for coupled atmosphere-snow/ice-ocean systems: A review and description of the AccuRT model. *Appl Sci*. 2018;8(12):2682.
50. IOCCG. Evaluation of atmospheric correction algorithms over turbid waters. In: Jamet C, editor. *Reports of International Ocean-Colour Coordinating Group*. Dartmouth (Canada): IOCCG; 2024. p. 112.
51. Hu C, Lee Z, Franz B. Chlorophyll aalgorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J Geophys Res Oceans*. 2012;117(C1):Article 2011JC007395.
52. Wang M, Liu X, Tan L, Jiang L, Son S, Shi W, Rausch K, Voss K. Impacts of VIIRS SDR performance on ocean color products. *J Geophys Res Atmos*. 2013;118(18):10347–10360.
53. Mikelsons K, Wang M. Characterization and removal of striping artifacts in VIIRS-derived ocean color products. *Opt Express*. 2025;33(3):5382–5395.
54. Zibordi G, D'Alimonte D, Kajiyama T. Automated quality control of AERONET-OC LWN data. *J Atmos Ocean Technol*. 2022;39(12):1961–1972.
55. Zibordi G, Mélin F, Berthon J-F, Holben B, Slutsker I, Giles D, D'Alimonte D, Vandemark D, Feng H, Schuster G, et al. AERONET-OC: A network for the validation of ocean color primary products. *J Atmos Ocean Technol*. 2009;26(8):1634–1651.
56. Lee Z, Carder KL, Mobley CD, Steward RG, Patch JS. Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Appl Opt*. 1999;38(18):3831–3843.
57. Mélin F, Zibordi G, Berthon J-F, Bailey S, Franz B, Voss K, Flora S, Grant M. Assessment of MERIS reflectance data as processed with SeaDAS over the European seas. *Opt Express*. 2011;19(25):25657–25671.
58. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323(6088):533–536.
59. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. Paper presented at: The 30th International Conference on Machine Learning; 2013 Jun 16–21; Atlanta, GA, USA.
60. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Paper presented at: The 27th international conference on machine learning; 2010 Jun 21–24; Haifa, Israel.
61. Safran I, Shamir O. Depth-width tradeoffs in approximating natural functions with neural networks. Paper presented at: The 34th International Conference on Machine Learning; 2017 Aug 6–11; Sydney NSW, Australia.
62. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov R. Improving neural networks by preventing co-adaptation of feature detectors. arXiv. 2012. https://doi.org/10.48550/arXiv.1207.0580
63. Shi W, Wang M. An assessment of the black ocean pixel assumption for MODIS SWIR bands. *Remote Sens Environ*. 2009;113(8):1587–1597.
64. Werdell PJ, Bailey SW. An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens Environ*. 2005;98(1):122–140.
65. Werdell PJ, Bailey S, Fargion G, Pietras C, Knobelspiesse K, Feldman G, McClain C. Unique data repository facilitates ocean color satellite validation. *EOS Trans Am Geophys Union*. 2003;84(38):377–387.
66. Lee Z, Pahlevan N, Ahn Y-H, Greb S, O'Donnell D. Robust approach to directly measuring water-leaving radiance in the field. *Appl Opt*. 2013;52(8):1693–1701.
67. Yu X, Lee Z, Shang Z, Lin H, Lin G. A simple and robust shade correction scheme for remote sensing reflectance obtained by the skylight-blocked approach. *Opt Express*. 2021;29(1):470–486.
68. Mueller JL, Morel A, Frouin R, Davis C, Arnone R, Carder K, Lee ZP, Steward RG, Hooker S, Holben B, et al. Radiometric measurements and data analysis protocols. In: Mueller JL, Fargion GS, McClain CR, editors. *Ocean optics protocols for satellite ocean color sensor validation, revision 4*. Greenbelt (MD): NASA; 2003.
69. Zibordi G, Voss KJ, Johnson BC, Mueller JL. Protocols for satellite ocean colour data validation: In situ optical radiometry. In: Zibordi G, editor*IOCCG ocean optics and*

*biogeochemistry protocols for satellite ocean colour sensor validation*Dartmouth (Canada): IOCCG; 2019. p. 67.
70. Ruddick KG, Voss K, Boss E, Castagna A, Frouin R, Gilerson A, Hieronymi M, Johnson BC, Kuusk J, Lee Z, et al. A review of protocols for fiducial reference measurements of water-leaving radiance for validation of satellite remote-sensing data over water. *Remote Sens*. 2019;11(19): Article 2198.
71. Lee Z, Wei J, Shang Z, Garcia R, Dierssen H, Ishizaka J, Castagna A. On-water radiometry measurements: Skylight-blocked approach and data processing. In: Zibordi G, Voss KJ, Johnson BC, Mueller JL, editors. *Protocols for satellite ocean colour data validation: In situ optical radiometry*. Dartmouth (Canada): IOCCG; 2019. p. 8.
72. Dong Q, Shang S, Lee Z. An algorithm to retrieve absorption coefficient of chromophoric dissolved organic matter from ocean color. *Remote Sens Environ*. 2013;128:259–267.
73. Shang S, Lee Z, Lin G, Hu C, Shi L, Zhang Y, Li X, Wu J, Yan J. Sensing an intense phytoplankton bloom in the western Taiwan Strait from radiometric measurements on a UAV. *Remote Sens Environ*. 2017;198:85–94.
74. Zhang J, Chen J, Xia X, Che H, Fan X, Xie Y, Han Z, Chen H, Lu D. Heavy aerosol loading over the Bohai Bay as revealed by ground and satellite remote sensing. *Atmos Environ*. 2016;124:252–261.
75. Li L, Bi X, Wang X, Song L, Dai Q, Liu B, Wu J, Zhang Y, Feng Y. High aerosol loading over the Bohai Sea: Long-term trend, potential sources, and impacts on surrounding cities. *Environ Int*. 2024;183:Article 108387.
76. Moulin C, Gordon HR, Chomko RM, Banzon VF, Evans RH. Atmospheric correction of ocean color imagery through thick layers of Saharan dust. *Geophys Res Lett*. 2001;28(1):5–8.
77. Shi W, Wang M. Detection of turbid waters and absorbing aerosols for the MODIS Ocean color data processing. *Remote Sens Environ*. 2007;110(2):149–161.
78. Al Shehhi MR, Gherboudj I, Zhao J, Ghedira H. Improved atmospheric correction and chlorophyll-a remote sensing models for turbid waters in a dusty environment. *ISPRS J Photogramm Remote Sens*. 2017;133:46–60.
79. Wang Y, Chen H-H, Tang R, He D, Lee Z, Xue H, Wells M, Boss E, Chai F. Australian fire nourishes ocean phytoplankton bloom. *Sci Total Environ*. 2022;807:Article 150775.
80. Geng X, Angal A, Sun J, Wu A, Choi T, Xiong X. Characterization of MODIS mirror side difference in the reflective solar spectral region. In: *Earth Observing Systems XVI, San Diego, California, United States, 21–25 August, 2011*. SPIE; 2011.
81. Mikelsons K, Wang M, Jiang L, Bouali M. Destriping algorithm for improved satellite-derived ocean color product imagery. *Opt Express*. 2014;22(23):28058–28070.
82. Kahn RA, Sayer AM, Ahmad Z, Franz BA. The sensitivity of SeaWiFS Ocean color retrievals to aerosol amount and type. *J Atmos Ocean Technol*. 2016;33(6):1185–1209.
83. Song Z, He X, Bai Y, Wang D, Gong F, Zhu Q, Li T, Li H. Effect of the vertical distribution of absorbing aerosols on the atmospheric correction for Satellite Ocean color remote sensing. *IEEE Trans Geosci Remote Sens*. 2022;60:1–12.
84. Mielonen T, Arola A, Komppula M, Kukkonen J, Koskinen J, de Leeuw G, Lehtinen KEJ. Comparison of CALIOP level 2 aerosol subtypes to aerosol types derived from AERONET inversion data. *Geophys Res Lett*. 2009;36(18):Article 2009GL039609.
85. Lee J, Kim J, Song CH, Kim SB, Chun Y, Sohn BJ, Holben BN. Characteristics of aerosol types from AERONET sunphotometer measurements. *Atmos Environ*. 2010;44(26):3110–3117.
86. Russell PB, Kacenelenbogen M, Livingston JM, Hasekamp OP, Burton SP, Schuster GL, Johnson MS, Knobelspiesse KD, Redemann J, Ramachandran S, et al. A multiparameter aerosol classification method and its application to retrievals from spaceborne polarimetry. *J Geophys Res Atmos*. 2014;119(16):9838–9863.
87. Bilal M, Ali MA, Nichol JE, Bleiweiss MP, de Leeuw G, Mhawish A, Shi Y, Mazhar U, Mehmood T, Kim J, et al. AEROsol generic classification using a novel satellite remote sensing approach (AEROSA). *Front Environ Sci*. 2022;10:Article 981522.
88. Holben BN, Eck TF, Slutsker I, Tanré D, Buis JP, Setzer A, Vermote E, Reagan JA, Kaufman YJ, Nakajima T, et al. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens Environ*. 1998;66(1):1–16.
89. Sinyuk A, Holben BN, Eck TF, Giles DM, Slutsker I, Korkin S, Schafer JS, Smirnov A, Sorokin M, Lyapustin A. The AERONET version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to version 2. *Atmos Meas Tech*. 2020;13(6):3375–3411.
90. Di Biagio C, Formenti P, Balkanski Y, Caponi L, Cazaunau M, Pangui E, Journet E, Nowak S, Andreae MO, Kandler K, et al. Complex refractive indices and single-scattering albedo of global dust aerosols in the shortwave spectrum and relationship to size and iron content. *Atmos Chem Phys*. 2019;19(24):15503–15531.
91. Kaskaoutis DG, Grivas G, Stavroulas I, Liakakou E, Dumka UC, Gerasopoulos E, Mihalopoulos N. Effect of aerosol types from various sources at an urban location on spectral curvature of scattering and absorption coefficients. *Atmos Res*. 2021;264:Article 105865.
92. Wei J, Lee Z, Shang S. A system to measure the data quality of spectral remote sensing reflectance of aquatic environments. *J Geophys Res Oceans*. 2016;121(11): 8189–8207.
93. Xiang H, Liu J. Detection and removal of the stripes in MODIS 1B image. *J Remote Sens*. 2011;15(1):123–136.
94. Mélin F, Sclep G. Band shifting for ocean color multi-spectral reflectance data. *Opt Express*. 2015;23(3):2262–2279.
95. Talone M, Zibordi G, Pitarch J. On the application of AERONET-OC multispectral data to assess satellite-derived hyperspectral R_{rs}. *IEEE Geosci Remote Sens Lett*. 2024;21:1–5.
96. Zibordi G, Mélin F, Berthon JF, Talone M. In situ autonomous optical radiometry measurements for satellite ocean color validation in the Western Black Sea. *Ocean Sci*. 2015;11(2):275–286.
97. Gordon HR, Brown OB, Evans RH, Brown JW, Smith RC, Baker KS, Clark DK. A semianalytic radiance model of ocean color. *J Geophys Res Atmos*. 1998;93(D9): 10909–10924.
98. Lee Z, Carder KL, Arnone RA. Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Appl Opt*. 2002;41(27):5755–5772.
99. Lee Z, Shang S, Lin G, Chen J, Doxaran D. On the modeling of hyperspectral remote-sensing reflectance of high-

sediment-load waters in the visible to shortwave-infrared domain. *Appl Opt*. 2016;55(7):1738–1750.
100. Lee Z, Wei J, Voss K, Lewis M, Bricaud A, Huot Y. Hyperspectral absorption coefficient of "pure" seawater in the range of 350–550 nm inverted from remote sensing reflectance. *Appl Opt*. 2015;54(3):546–558.
101. Zhang X, Hu L, He M-X. Scattering by pure seawater: Effect of salinity. *Opt Express*. 2009;17(7):5698–5710.
102. Lee Z, Ahn Y-H, Mobley C, Arnone R. Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform. *Opt Express*. 2010;18(25):26313–26324.
103. Werdell PJ, Franz BA, Bailey SW, Feldman GC, Boss E, Brando VE, Dowell M, Hirata T, Lavender SJ, Lee Z, et al. Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl Opt*. 2013;52(10):2019–2037.