



An improved spectral optimization algorithm for atmospheric correction over turbid coastal waters: A case study from the Changjiang (Yangtze) estuary and the adjacent coast



Yanqun Pan^a, Fang Shen^{a,*}, Wouter Verhoef^b

^a State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200062, China

^b Department of Water Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands

ARTICLE INFO

Article history:

Received 13 April 2016

Received in revised form 23 December 2016

Accepted 14 January 2017

Available online xxxx

Keywords:

Ocean color

Atmospheric correction

GOCI

MERIS

Turbid coastal waters

ABSTRACT

Remote sensing-based retrieval of the concentrations of water components relies largely on the accuracy of the atmospheric correction. Although a variety of atmospheric correction algorithms have been developed for turbid waters, the water-leaving reflectance is still underestimated in extremely turbid waters, such as in the Changjiang (Yangtze) estuary and the adjacent coast. To address this issue, this paper proposes an improved algorithm that is based on a spectral optimization algorithm (ESOA) with a coupled water-atmosphere model. The model combines an aerosol model that is constructed from Aerosol Robotic Network (AERONET) observation data and a simple semi-empirical radiative transfer (SERT) model (Shen et al. 2010) for water component retrieval. Four unknown parameters are involved in the coupled model: the relative humidity (RH), fine-mode fraction (FMF), aerosol optical thickness in the near-infrared (NIR) wavelength $\tau_a(\lambda_0)$ and suspended particulate matter (SPM) concentration (C_{spm}). These parameters are estimated by a global optimization approach that is based on a genetic algorithm (GA) without any initial inputs. Validation results of the atmospherically corrected remote sensing reflectance $R_{rs}(\lambda)$ from matchups between Geostationary Ocean Color Imager (GOCI) data and in situ data show that the algorithm has satisfactory accuracy. The root mean square error (RMSE) and the absolute percentage difference (APD) are 0.0089 and 35.12, respectively. By contrast, the $R_{rs}(\lambda)$ values retrieved from the same matchups using the GOCI data processing system (GDPS) have higher RMSE and APD of 0.0104 and 69.15, respectively. The ESOA method can be implemented conveniently within the open source code of SeaDAS (v7.1) as an alternative and operational tool for atmospheric correction of ocean color data, including GOCI, MERIS and MODIS, over highly turbid estuarine and coastal regions, such as the Yangtze estuary, the Hangzhou Bay and most of the coastal ocean in Eastern China.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Approximately 90% of the radiant signals received by ocean color remote sensors is contributed by atmospheric components that are unrelated to the water components. Therefore, the accuracy of atmospheric correction critically affects the accuracy of ocean color parameter retrievals. In recent decades, due to the efforts of many researchers, the atmospheric correction for Case-1 waters has been well established. However, the atmospheric correction for turbid coastal waters is still unsatisfactory. The difficulty in atmospheric correction is how to determine the types and optical thicknesses (τ_a) of aerosols based on remote-sensing images. A classical atmospheric correction algorithm (GW94; Gordon and Wang, 1994) that is based on the near-infrared (NIR) “dark pixel” assumption infers the ratio (ε) of the aerosol scattering

reflectances ρ_a at two NIR bands ρ_a (NIR). The algorithm neglects the absorption of heavily absorbing aerosols, selects the best aerosol model from 12 candidate aerosol models (M50, M70, M90, M99, C50, C70, C90, C99, T50, T80, T99, and O99) based on ε and then extrapolates ρ_a (NIR) to the visible spectrum (VIS). In the implementation of GW94 in the Sea-viewing Wide Field-of-View Sensor Data Analysis System (SeaDAS) (v7.1), the 12 candidate aerosol models are replaced with the AERONET-based aerosol model (Ahmad et al., 2010), which works well for Case-1 waters but suffers from two problems when applied to turbid waters. First, the assumption of an NIR “dark pixel” is invalid. Second, the absorption by heavily absorbing aerosols cannot be ignored because most turbid waters are located in estuaries and coastal regions. Due to the influence of continental contaminant emissions (e.g., smoke), aerosols over these regions are often absorbing. Numerous researchers have attempted to address these two problems.

For example, using iterative schemes, Siegel et al. (2000) and Bailey et al. (2010) estimated chlorophyll-*a* (Chl_a) concentrations by

* Corresponding author.

E-mail address: fshen@sklec.ecnu.edu.cn (F. Shen).

establishing a semi-analytical model of the water-leaving radiance L_w and the Chla concentration. Shi et al. (2012) derived the diffuse attenuation coefficient at 490 nm (K_d_{490}) by developing an empirical relationship between L_w and the water diffuse attenuation coefficient K_d , deriving L_w in the NIR using a bio-optical model, and gradually decreasing L_w in the NIR using an iterative scheme. Atmospheric correction algorithms that are based on iterative schemes have improved the accuracy of atmospheric correction for turbid waters. However, these iterative schemes are fundamentally based on the NIR “dark pixel” assumption, which overestimates $\rho_a(\lambda)$ at the beginning of the iteration, which will likely result in negative L_w values in the blue region. To address the issue that the L_w values of turbid waters for short-wavelength infrared (SWIR) spectra are close to 0, Wang and Shi (2005) and Wang et al., 2009 proposed the assumption of a SWIR “dark pixel” for MODIS atmospheric correction. This method has been validated over western Pacific turbid waters. However, most current ocean color remote sensors, such as the Sea-viewing Wide Field-of-View Sensor (SeaWiFS), Medium Resolution Imaging Spectrometer (MERIS) and Geostationary Ocean Color Imager (GOCI), do not include SWIR bands, which limits the application of this method. He et al. (2012) proposed an atmospheric correction algorithm based on the assumption of an ultraviolet (UV) “dark pixel” (UV-AC). The UV-AC algorithm produces good corrections for turbid waters, but the UV “dark pixel” assumption is still invalid for highly turbid waters (He et al., 2012; Knaeps et al., 2012). In addition to the “dark pixel” assumption, Ruddick et al. (2000) proposed an atmospheric correction algorithm that assumes that the ratios between ($\rho_a(765)/\rho_a(865)$, $\rho_w(765)/\rho_w(865)$) are homogenous over a study area. Due to the large temporal and spatial variabilities of turbid estuaries and coastal regions (for example, the Yangtze estuary and Hangzhou Bay), the assumption of a homogenous ρ_w ratio is always invalid. Mao et al. (2013, 2014) suggested that using ε between two infrared bands to extrapolate ρ_a (VIS) would amplify the errors. Thus, they proposed a new method to determine the aerosol type by matching measured $R_{rs}(\lambda)$ data with aerosol models. This method requires a large amount of measurements to build a remote-sensing reflectance database. On the one hand, if the volume of measured data is not sufficient, it is difficult to fully represent all of the water's spectral characteristics; on the other hand, if the volume of measured data is too large, it will take a long time to perform searches, which reduces the practicality of the algorithm.

The atmospheric correction algorithms described above are all based on the 12 classical types of aerosol models, which use ε to determine the aerosol types and extrapolate ρ_a (VIS). This method is highly reliable for non-absorbing or weakly absorbing aerosol models but causes large errors when considering the absorption of heavy-absorbing aerosols (Gordon et al., 1997). Gordon et al. (1997) proposed an atmospheric correction algorithm based on spectral matching (SMA), and Chomko and Gordon (1998) proposed an atmospheric correction algorithm based on spectral optimization (SOA). These two algorithms build a coupled water-atmosphere model that simultaneously establishes $\rho_w(\lambda)$ and the aerosol type. By finding the optimal combination of the water spectral reflectance and aerosol reflectance, these methods simultaneously derive the water's bio-optical and aerosol model parameters. The difference between the methods is that the former searches the discrete aerosol models one-by-one, whereas the latter uses the traditional nonlinear optimization method. The SOA model assumes that a simple single-parameter model represents the particle size distribution of the aerosol, uses a series of complex refractive indices to represent the aerosol absorptivity within a certain range and then calculates the aerosol optical properties using Mie scattering theory. This method also selects a semi-analytical bio-optical model based on the Chla concentration C_{phy} and the particle scattering coefficient b_0 as the water-leaving radiance model. This coupled model includes six parameters (m_r , m_i , ν , τ , C_{phy} , and b_0) to be optimized. With the NIR “dark pixel” assumption, these are reduced to four parameters (ν , τ , C_{phy} , and b_0). Lastly, it uses a constrained nonlinear optimization method to

simultaneously retrieve the four parameters. Chomko and Gordon (1998) applied the SOA algorithm to the atmospheric correction of SeaWiFS imagery over open ocean waters and validated the retrieved aerosol parameters and water bio-optical parameters. Chomko et al. (2003) improved the SOA algorithm by combining it with a globally tuned version of the Garver and Siegel (1997) bio-optical model (GSM01) and conducting an initial estimation of parameters using the NIR “dark pixel” assumption. Because the NIR “dark pixel” assumption is invalid for Case-2 waters, Kuchinke et al. (2009) used an iterative method to make an initial estimation of the parameters based on the NIR “dark pixel” assumption and proposed an improved SOA algorithm (SOA2009) so it could be applied to the atmospheric correction for Case-2 waters. Comparisons between the correction and modeling results and measurements in Chesapeake Bay showed that the algorithm performed very well.

Although SOA2009 provides a good model for atmospheric correction over turbid waters, its application in regional waters with complex optical properties, such as the Yangtze estuary and the adjacent coast, is limited for several reasons. (1) The particle size distribution of the aerosol model in SOA2009, which is based on a Junge power law (see Chomko and Gordon (1998) and Kuchinke et al. (2009) for a description), cannot explain the appearance of large particles in the observed particle size distribution (Davies, 1974), especially in coastal regions. (2) The GSM01-based SOA2009 relies on the chlorophyll-specific absorption spectra a_{ph} , the colored dissolved organic matter (CDOM) spectral slope S and the particle backscattering spectra n , which vary strongly in waters with regionally complex optical properties. In addition, in highly turbid waters, L_w is mainly determined by the backscattering of suspended particulate matter (SPM). The GSM01 model, which synthesizes the effects of Chla, CDOM and total suspended matter (TSM), is too complicated. (3) SOA2009 uses a traditional constrained nonlinear optimization algorithm, such as the quasi-Newton algorithm, that strongly relies on parameter initialization and can only find an optimum near the initial values. However, it is difficult to accurately estimate the initial values of the parameters.

This study proposes an improved SOA atmospheric correction algorithm for turbid waters (ESOA). To address the three limitations described above, we made the following improvements: (1) In ESOA, the aerosol model is based on the AERONET observation data; thus, it can accurately reflect the actual conditions of the coastal aerosols (Ahmad et al., 2010). Additional details about this model are given in Section 2.1. (2) ESOA replaces the GSM01 model with a simple semi-empirical radiative transfer (SERT) model that has fewer parameters and works well with turbid coastal waters (Section 2.2). By combining (1) and (2), we derive a set of nonlinear equations based on radiative transfer (Section 3). (3) ESOA replaces the traditional optimization methods with a global-optimization genetic algorithm that does not rely on parameter initialization (Section 3). In Section 4, we validate ESOA with the simulated data and measurements separately. The measurements that are used for the validation include measured $R_{rs}(\lambda)$ data, fixed station SPM datasets and GOCI images over the Yangtze estuary and the adjacent coast. Finally, this study discusses the operational satellite image processing approach based on ESOA.

2. Aerosol and water models

In this study, we replace the top of atmosphere (TOA) radiance L with the planetary reflectance, which is defined as:

$$\rho = \frac{\pi L}{F_0 \cos \theta_0}, \quad (1)$$

where F_0 is the extraterrestrial solar irradiance, and θ_0 is the solar zenith angle. Thus, the TOA reflectance of the ocean-atmosphere system at wavelength λ is $\rho_t(\lambda)$, and $\rho_m(\lambda)$ is the calibrated $\rho_t(\lambda)$ with Rayleigh scattering $\rho_r(\lambda)$ correction, white cap reflectance $\rho_{wc}(\lambda)$ and flare

reflection $\rho_g(\lambda)$ using the SeaDAS (v7.1 <http://seadas.gsfc.nasa.gov/>) standard algorithm (Eq. (2)),

$$\rho_m(\lambda) = [\rho_t(\lambda) - \rho_r(\lambda) - t(\lambda)\rho_{wc}(\lambda) - T(\lambda)\rho_g(\lambda)], \quad (2)$$

where $t(\lambda)$ and $T(\lambda)$ are the diffuse and direct transmittances in the sea surface-to-sensor direction, respectively. In the SeaDAS standard algorithm, the calculation of $\rho_r(\lambda)$ is based on the model proposed by Gordon et al. (1988), $\rho_{wc}(\lambda)$ is estimated using an improved model in which information on wind-wave conditions were taken into account (Stramska and Petelski, 2003), and $\rho_g(\lambda)$ is estimated using an azimuthally symmetric form of Cox-Munk model (Wang and Bailey, 2001). The sum of $\rho_a(\lambda)$ and $\rho_w(\lambda)t(\lambda)$ is represented as:

$$\rho_{aw}(\lambda) = \rho_a(\lambda) + \rho_w(\lambda)t(\lambda). \quad (3)$$

The terms on the right-handside of Eq. (3) depend on the specific aerosol and water models. In theory, $\rho_m(\lambda)$ is equal to $\rho_{aw}(\lambda)$ at all wavelengths. Thus, the optimization algorithm is used to estimate the unknown parameters. The following section fully describes the aerosol model, the water model and their parameters that were used in this study.

2.1. The aerosol model

Because aerosol models with power-law particle-size distributions cannot accurately explain the appearance of large particles in observed particle-size distributions, Davies proposed a lognormal function for aerosol particle-size distributions (Davies, 1974). Shettle and Fenn (1979) classified the sources of aerosols. Based on tropospheric and maritime type aerosols, Gordon and Wang (1994) proposed a series of aerosol models for the atmospheric correction of remote-sensing imagery. However, these models cannot be used for smoke and dust over the ocean (Ahmad et al., 2010). Ahmad et al. (2010) analyzed approximately ten years of data from eight AERONET stations (three coastal stations and five marine stations) and discovered that the seasonal variability in the aerosol radius was correlated to the relative humidity. Based on that finding, he developed an aerosol model for the inversion of remotely sensed atmospheric optical properties (referred to as Ahmad2010). Ahmad2010 includes eight coarse-mode particles that are represented by RH (30%, 50%, 70%, 75%, 80%, 85%, 90%, and 95%). Every coarse mode is composed of 10 fine-mode (0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.5, 0.8, and 0.95) particles that are represented by fine-mode fractions (FMFs). A total of 80 types of aerosols were created. Ahmad et al. (2010) assumes that the coarse-mode particles represent non-absorbing aerosols such as sea salt; thus, the absorption observed by AERONET is entirely due to fine-mode particles. Ahmad used Ahmad2010 to replace the 12 classic types of aerosol models for atmospheric correction, and the derived τ_a and L_w values were more accurate. Thus, ESOA uses Ahmad2010 as the aerosol model. The look-up tables (LUTs) are included in SeaDAS. Each LUT includes optical parameters, such as the aerosol extinction coefficient $c(\lambda)$, single scattering albedo $\omega_0(\lambda)$, phase function $P(\theta, \theta_0, \lambda)$, Ångström coefficient, quadratic coefficients of the single-scattering to multi-scattering function, coefficients of diffuse sensor transmittance, and coefficients of diffuse solar transmittance. In the SeaDAS standard algorithm for the LUTs, aerosol optical properties such as $\rho_a(\lambda)$, $t(\lambda)$ and $\tau_a(\lambda)$ are calculated for a given $\tau_a(\lambda_0)$. The detailed calculation process is found in the atmospheric correction module in SeaDAS which is open source (<https://oceancolor.gsfc.nasa.gov/ocssw>), and also is described by Mobley et al. (2016). Because these are limited to the 80 LUT entries for the Ahmad2000 aerosol model that are represented by eight RH values and ten FMF values, bilinear interpolation is used to calculate $\tau_a(\lambda)$, $\rho_a(\lambda)$ and $t(\lambda)$ for aerosol types that are represented by arbitrary values of RH and FMF. Fig. 1(a) shows the aerosol reflectance $\rho_a(\lambda)$, and Fig. 1(b) shows the diffuse

transmittance $t(\lambda)$ in different aerosol models for an RH of 70% and a solar zenith angle, sensor zenith angle, relative angle, and $\tau_a(865)$ of 30°, 15°, 120°, and 0.1, respectively, where λ is the central wavelength of each GOCI band.

2.2. The semi-empirical radiative transfer (SERT) model

The remote-sensing reflectance is related to the inherent optical properties (IOPs) of the water body via the underwater reflectance (r). Kubelka and Munk (KM) (1931) proposed a calculation model for r from a semi-infinite turbid medium based on a two-stream approximation in the radiative transfer:

$$r = \frac{b_b/a}{b_b/a + 1 + \sqrt{2b_b/a}}. \quad (4)$$

Eq. (4) shows that r is only a function of the ratio $x = b_b/a$. For highly turbid waters, Shen et al. (2010a,b) assumed that x was positively correlated with the SPM concentration and proposed a SERT model based on that assumption:

$$R_{rs}(\lambda) = \frac{\alpha(\lambda)\beta(\lambda)C_{spm}}{1 + \beta(\lambda)C_{spm} + \sqrt{1 + 2\beta(\lambda)C_{spm}}}, \quad (5)$$

where α and β are model coefficients that are wavelength-dependent, and C_{spm} is the SPM concentration. For waters in which the SPM is the dominant variable, one may assume that x is a function of C_{spm} as follows:

$$x = \frac{b_{bO} + C_{spm}b_{bS}}{a_O + C_{spm}a_S}, \quad (6)$$

where the subscript S refers to suspended sediment, and O stands for all of the other constituents of the water, including the pure water itself. Neglecting the absorption of SPM and the backscatter of the other water constituents, x becomes proportional to C_{spm} ($x = \beta C_{spm}$). The other constituents are the pure water itself and the (assumed constant) amounts of CDOM and Chla. To compensate R_{rs} for the influence of the absorption of SPM, another parameter α is introduced in the SERT model. The values of α and β in the SERT model at different wavelengths were fitted from measured R_{rs} and SPM concentrations. This means that the amounts of CDOM and Chla are assumed to be constant in the SERT model, so the influences of their variations are neglected.

Shen et al. (2013, 2014) estimated the SPM concentrations in the Yangtze estuary and the adjacent coast through the application of the SERT model to numerous types of ocean color sensors (e.g., GOCI, MERIS and MODIS) and analyzed the long-term time series of SPM concentrations. After comparison and validation with multiple measurements, they found that the SPM concentrations derived from three sensors with nearly synchronous overpasses (MERIS, MODIS, and FY-3 MERSI) exhibited excellent consistency, with mean differences of 0.056, 0.057, and 0.013 g l⁻¹, respectively, at three field fixed stations in the Yangtze estuary. By combining Eq. (5), the definition of $\rho_w = L_w / F_0 \cos(\theta_0)$ in this study (based on Eq. (1)) and $R_{rs} = L_w / E_d(0+)$, where $E_d(0+)$ refers to the downwelling irradiance just above the surface which is approximately equal to $F_0 \cos(\theta_0)t_0$, one can derive the following equation:

$$\rho_w(\lambda) = \frac{\alpha(\lambda)\beta(\lambda)C_{spm}}{1 + \beta(\lambda)C_{spm} + \sqrt{1 + 2\beta(\lambda)C_{spm}}} \pi t_0(\lambda), \quad (7)$$

where $t_0(\lambda)$ is the diffuse transmittance in the sun-to-sea direction.

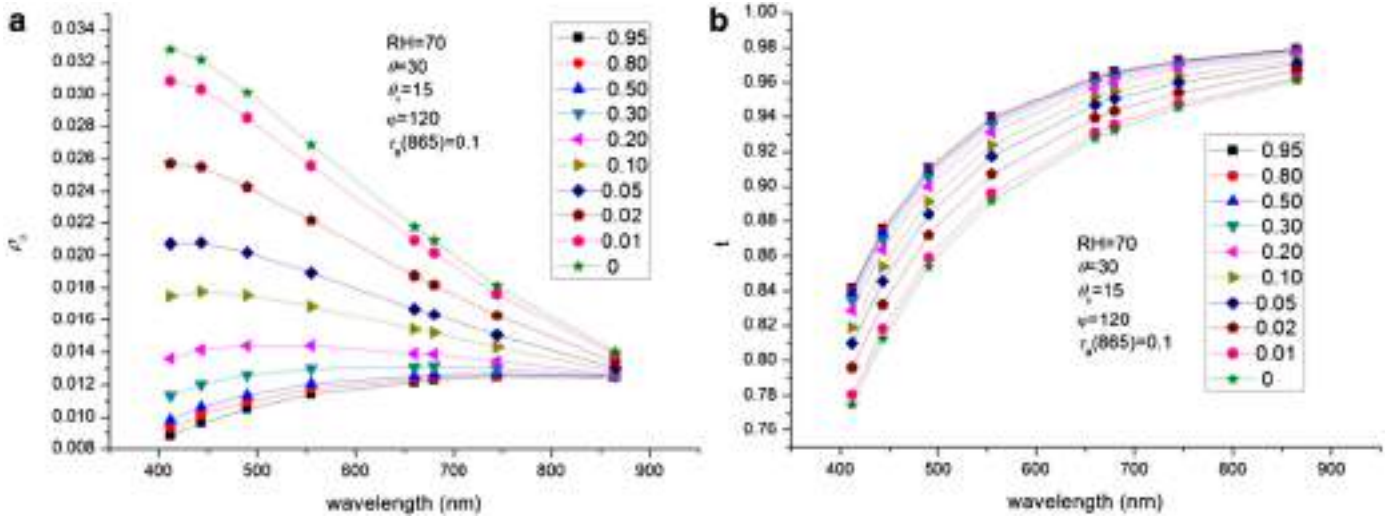


Fig. 1. Aerosol reflectance $\rho_a(\lambda)$ (a) and diffuse transmittance from the sea surface to the sensor $t(\lambda)$ (b) calculated by the aerosol model developed by Ahmad for different FMFs (upper to lower curves) for RH = 70%, solar zenith angle $\theta_0 = 15^\circ$, sensor zenith angle $\theta = 30^\circ$, relative azimuth angle $\varphi = 120^\circ$ and $\tau_a(865) = 0.1$.

3. Implementation of ESOA based on a genetic algorithm

In Section 2, we described the basic process of radiative transfer, the Ahmad2010 aerosol model and the SERT water model. Assuming that the four unknown parameters (RH, FMF, $\tau_a(\lambda_0)$ and C_{spm}) in the water-atmosphere model are known, then $\rho_a(\lambda)$, $t_0(\lambda)$, $t(\lambda)$ and $\rho_w(\lambda)$ can be estimated from these four parameters, which are denoted as $\rho_a(\lambda, \text{RH, FMF, } \tau_a(\lambda_0))$, $t_0(\lambda, \text{RH, FMF, } \tau_a(\lambda_0))$, $t(\lambda, \text{RH, FMF, } \tau_a(\lambda_0))$ and $\rho_w(\lambda, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm})$, respectively. Thus, $\rho_{aw}(\lambda, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm})$ can be obtained by Eq. (8):

$$\rho_{aw}(\lambda, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm}) = \rho_a(\lambda, \text{RH, FMF, } \tau_a(\lambda_0)) + \rho_w(\lambda, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm})t(\lambda, \text{RH, FMF, } \tau_a(\lambda_0)). \quad (8)$$

By combining Eq. (2), Eq. (3) and the water-atmosphere model described above, one can build the following set of nonlinear equations for sensors with the waveband number N for the visible and near-infrared wavelengths:

$$\begin{aligned} \rho_m(\lambda_1) &= \rho_{aw}(\lambda_1, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm}), \\ \rho_m(\lambda_2) &= \rho_{aw}(\lambda_2, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm}), \\ \rho_m(\lambda_N) &= \rho_{aw}(\lambda_N, \text{RH, FMF, } \tau_a(\lambda_0), C_{spm}). \end{aligned} \quad (9)$$

SOA2009 and its prior SOA algorithms use traditional constrained nonlinear optimization algorithms, such as Broyden Fletcher Garbo Shannon (BFGS) (Byrd et al., 1995), for parameter optimization. However, this type of nonlinear optimization algorithm relies on variable initialization. They can only find the local optimum within the solution space set by the initial values but cannot find the global optimum. It is difficult to accurately estimate the initial values of the four unknown parameters for the nonlinear problem in Eq. (9). Thus, traditional nonlinear optimization algorithms are not applicable to this nonlinear problem if an accurate first estimate is not available. Genetic algorithms (GAs) are a type of global optimization algorithm that do not rely on differential and initial values of the variables. Thus, a GA is introduced to perform the optimization of Eq. (9). The genetic algorithm modeling for a specific problem generally needs to identify three important aspects: the ranges and encodings of the variables, the objective function (or fitness function) and the genetic operators and their related parameters. The details of the GA implementation for the specific nonlinear problem in this study are described in the Appendix.

4. Validation

We described the aerosol model and water model of ESOA and the genetic algorithm for spectral optimization in detail in the previous sections. To validate the performance of the algorithm, we established a simulated dataset and validated the algorithm via the optimization of the simulated data. We then established matchup datasets for in situ measurements and remote-sensing imagery. We validated the algorithm via a comparison between the in situ measurements and the retrieved results for the matchups. Due to limitations in weather conditions and the measurement cruises, few matchup pairs of $R_{rs}(\lambda)$ were available. Therefore, we also compared the SPM derived by the ESOA algorithm with the in situ SPM measurements.

4.1. Validation with simulated data

To validate the ESOA algorithm and determine its performance, ESOA was first tested with synthetic data that contained different levels of noise (0, 2%, and 5%). A similar method to construct simulated datasets was used by Maritorena et al. (2002). We created synthetic datasets using the aerosol and water models described above in a forward model with different values of the four parameters (RH, FMF, $\tau_a(865)$ and C_{spm}), which were randomly generated with ranges of [30,95], [0,100], [0.01,0.5] and [20,1000], respectively (see details of the ranges in the Appendix). The values of the observation geometry parameters (solar zenith angle θ_0 , viewing zenith angle θ and relative azimuth angle φ) were also generated randomly with ranges of [0,60], [0,60] and $[-180,180]$, respectively. We simulated 1000 samples of the $\rho_m(\lambda)$ spectrum with these random values using Eq. (8).

Two additional synthetic datasets with added noise were created. Random noise was added to the $\rho_m(\lambda)$ spectra by multiplying the original signal by a signal that consisted of spectrally uncorrelated normally distributed random deviations with a mean of 1.0 and a standard deviation of 0.02 or 0.05 (referred to as the 2% and 5% noise datasets, respectively). Whether this scheme reproduces realistic situations does not actually matter because these data are designed to assess ESOA and how it accommodates noisy data.

Table 1
Two parameters of the SERT model at the GOCI bands.

Waveband (nm)	412	443	490	555	660	680	745
α	0.0201	0.0253	0.0311	0.0488	0.0771	0.0797	0.0954
β	49.6982	48.3820	47.510	33.713	11.016	10.2475	2.9698

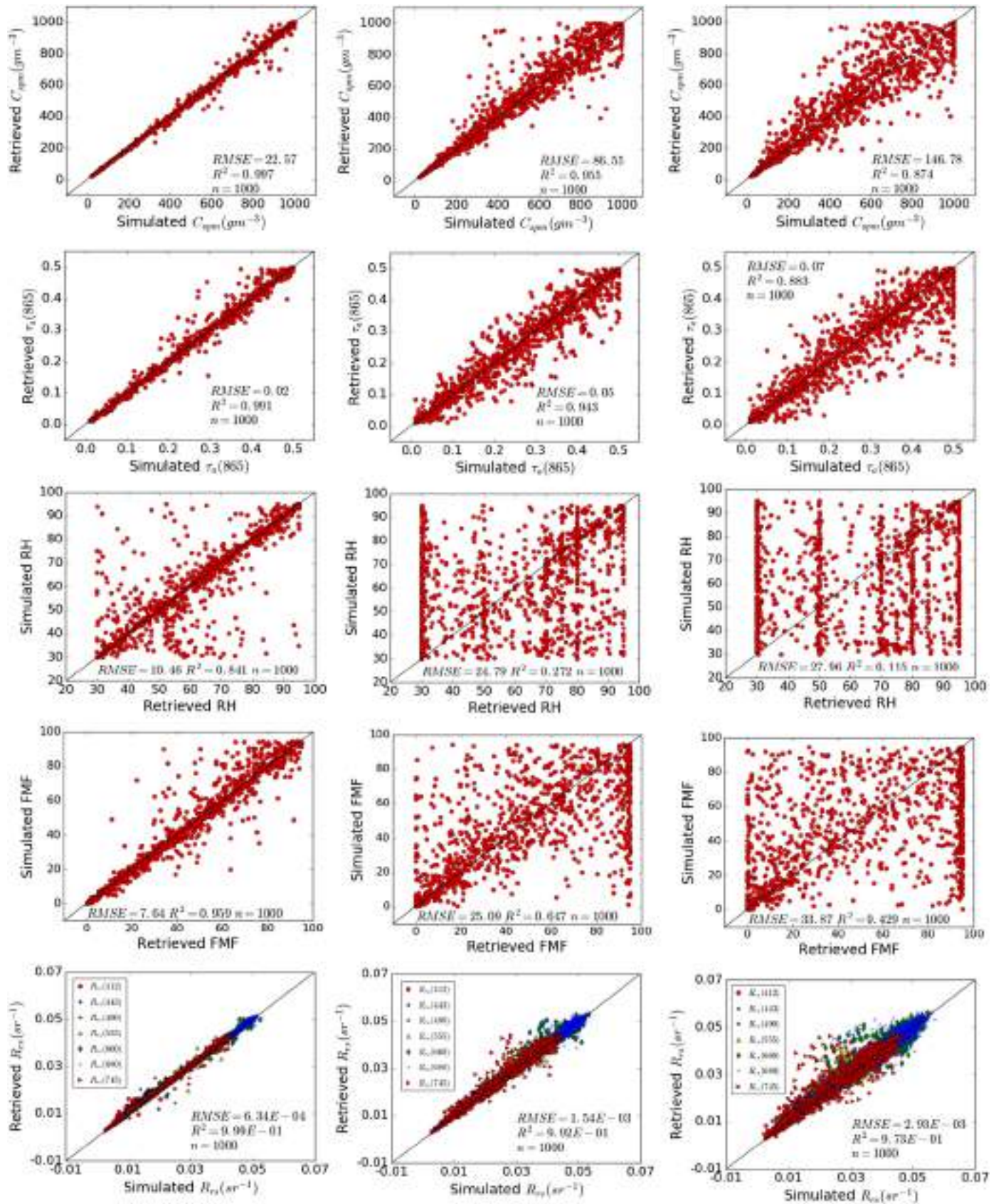


Fig. 2. Retrieved SPM, τ_{865} , RH, FMF and $R_{rs}(\lambda)$ values (from upper to lower) versus the actual values in synthetic datasets with no noise (left panels), 2% noise (middle panels), and 5% noise (right panels).

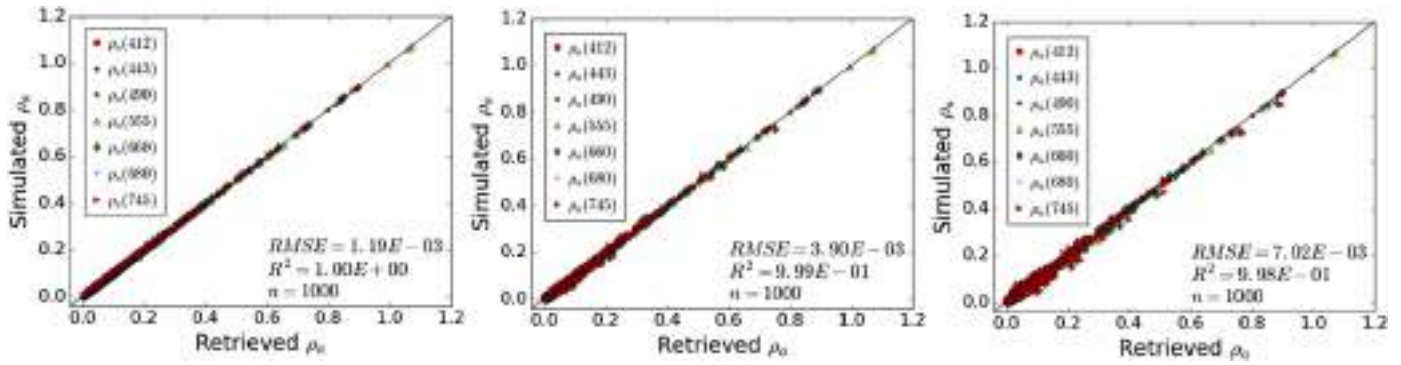


Fig. 3. Retrieved $\rho_a(\lambda)$ values versus actual values in the synthetic datasets with no noise (left), 2% noise (middle), and 5% noise (right).

The ESOA algorithm was first tested with the three synthetic datasets. All of the tests were conducted with the first seven GOCI bands. Table 1 shows the values of α and β in the SERT model for the GOCI bands. Because the longest wavelength of the R_{rs} measurements used to fit the model coefficients α and β was 858 nm, there are no values of α and β at 865 nm.

After optimizing the four variables, the $R_{rs}(\lambda)$ spectrum was estimated using the inverse model (Eq. (10)):

$$R_{rs}(\lambda) = \frac{\rho_m(\lambda) - \rho_a(\lambda, RH^0, FMF^0, \tau_a(865)^0)}{t_0(\lambda, RH^0, FMF^0, \tau_a(865)^0) t(\lambda, RH^0, FMF^0, \tau_a(865)^0) \pi}, \quad (10)$$

where RH^0 , FMF^0 and $\tau_a(865)^0$ refer to the RH, FMF and $\tau_a(865)$ values retrieved by ESOA, respectively.

For the dataset with no added noise, the retrieved variables are similar to the initial values. The RMSEs and correlation coefficients R^2 for the five variables are 22.57 and 0.997 (C_{spm}), 0.02 and 0.991 ($\tau_a(865)$), 10.46 and 0.841 (RH), 7.64 and 0.959 (FMF) and $6.34E-4$ and 0.99 ($R_{rs}(\lambda)$), respectively. These excellent agreements demonstrate that the approach can optimize the four unknowns in the complex, nonlinear system well. For datasets with random noise, the optimization errors for RH and FMF increase significantly and are almost entirely off the 1:1 line (Fig. 2). The main reason for the inconsistent

results for RH and FMF is that the set of nonlinear equations (Eq. (9)) has more than one optimum solution for the two parameters. This suggests that different combinations of RH and FMF may generate very similar $\rho_a(\lambda)$ spectra (Fig. 3). For the dataset without added noise, the RMSE and R^2 values between the simulated and retrieved $\rho_a(\lambda)$ values are $5.59E-4$ and 1.00, respectively. For the datasets with 2% and 5% noise added, the RMSE and R^2 values are $4.41E-3$ and 0.999 and $2.38E-3$ and 0.980, respectively. The retrieval errors for C_{spm} , $\tau_a(865)$ and $R_{rs}(\lambda)$ also increase with the amount of noise, but the results generally agree well with the measurements (Fig. 2). The modeled C_{spm} and $R_{rs}(\lambda)$ values were retrieved with high fidelity throughout the range of concentrations even with 5% noise. This demonstrates that ESOA can determine reasonably successful parameter candidates for C_{spm} , $\tau_a(865)$ and $R_{rs}(\lambda)$ even in the presence of significant noise. RH and FMF cannot be retrieved accurately by the proposed algorithm; however, this does not have a significant influence on the $R_{rs}(\lambda)$ retrieval.

To further evaluate the performance of the GA for this set of nonlinear equations (Eq. (9)), we also processed the synthetic datasets using the Levenberg-Marquardt (LM) algorithm (Kanzow et al., 2005) with random initial parameter values and compared the RMSEs of the parameters SPM, $\tau_a(865)$, $R_{rs}(\lambda)$ and $\rho_a(\lambda)$. Fig. 4 shows the RMSEs computed for the four parameters when the LM and GA are used to optimize the synthetic datasets with 0, 2%, and 5% noise. The histograms show that the GA performs better than the LM algorithm because,

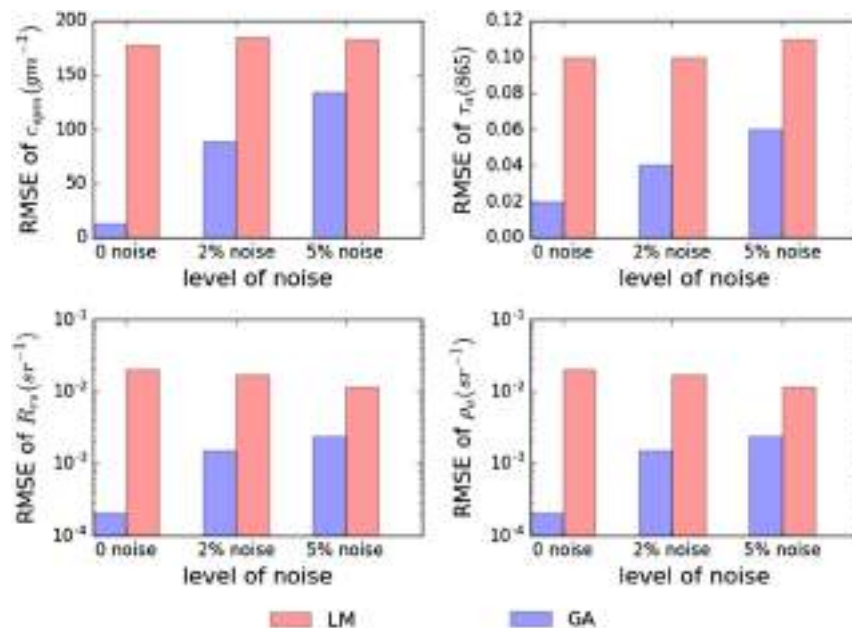


Fig. 4. RMSEs computed for SPM, $\tau_a(865)$, $R_{rs}(\lambda)$ and $\rho_a(\lambda)$ when the Levenberg-Marquardt algorithm (LM) and the genetic algorithm (GA) are used to optimize the synthetic datasets with 0, 2% and 5% noise.

Table 2

RMSE values of C_{spm} , $\tau_a(865)$, RH, FMF and R_{rs} for synthetic datasets with different observation geometries (θ , θ_0 and φ) without random noise.

$\theta, \theta_0, \varphi$	RMSE				
	C_{spm}	$\tau_a(865)$	RH	FMF	R_{rs}
5, 5, 30	22.76	0.03	7.29	14.45	0.000406
5, 5, 90	18.52	0.01	5.67	6	0.000301
5, 5, 150	23.73	0.02	4.7	6.38	0.000408
15, 15, 30	6.6	0.01	4.8	2.06	0.000101
15, 15, 90	5.48	0.03	10.98	7.42	8.19E-05
15, 15, 150	10.32	0.02	5.49	2.74	0.000153
30, 30, 30	6.86	0.01	3.16	1.66	9.22E-05
30, 30, 90	2.43	0.01	3.63	2.15	4.61E-05
30, 30, 150	8.74	0.03	6	6.72	0.000109
50, 50, 30	25.67	0.01	7.44	2.76	0.000349
50, 50, 90	5.67	0.03	14.97	6.99	8.76E-05
50, 50, 150	9.03	0.04	10.85	7.7	0.000167
70, 70, 30	59.14	0.01	3.11	1.59	0.000942
70, 70, 90	48.54	0.02	9.6	7.84	0.000616
70, 70, 150	11.09	0.01	4.95	5	0.000179

unlike LM, the GA does not depend on the parameter initialization. Because the first estimates for the parameters are random, the LM algorithm cannot find the global optimum for the synthetic datasets regardless of the amount of noise. The GA is more sensitive to noise than the LM algorithm. For example, there is no significant difference between the RMSEs computed by the LM algorithm with different noise levels for C_{spm} , while the RMSE for the GA increased with increasing amounts of noise. However, it is worth noting that the RMSE computed with a specific amount of noise is lower for the GA than for the LM algorithm; this is true for all of the parameters. The advantage of the GA increases with decreasing amounts of noise.

To discuss the sensitivity of ESOA to the observation geometry, we generated 15 groups of simulated datasets with five combinations of θ and θ_0 ((5, 5), (15, 15), (30, 30), (50, 50), and (70, 70)) and three values of φ (30, 60, and 90). For each group, the parameters other than the observation geometry (C_{spm} , $\tau_a(865)$, RH and FMF) that were used to

generate the simulated datasets are the same. Each group contains 1000 samples. These simulated datasets were optimized with ESOA. The results show that the observation geometry has a significant influence on the retrieval of C_{spm} , $\tau_a(865)$, RH, FMF and R_{rs} . Table 2 lists the RMSEs of each parameter for each group of datasets. The RMSE for C_{spm} has a significant variation from 2.43 to 59.14. For the same θ and θ_0 values, the lowest RMSE occurs when the relative azimuth φ is 90. In addition, for the same relative azimuth φ , the lowest RMSE occurs with solar and sensor zenith angles of 30 and 30, respectively. The parameter R_{rs} has a similar RMSE pattern as C_{spm} . However, there are no obvious patterns in the RMSE values for $\tau_a(865)$, RH and FMF.

4.2. Validation with matchup datasets of $R_{rs}(\lambda)$

We verified the ESOA algorithm with GOCI images and R_{rs} measurements taken during May 2011, March 2012 and March 2013 over the Yangtze River estuary and the adjacent seas. Radiometric measurements were recorded by the Hyperspectral Surface Acquisition System (Hyper SAS, Satlantic Inc., Canada) and include the sea-surface radiance (L_t), sky radiance (L_i) and solar irradiance (E_s) from 350 to 900 nm in 136 spectral channels. Details of the sensor mounting and observation geometries can be found in Shen et al. (2014). The values of R_{rs} were calculated by (Hooker et al., 2002; Mobley, 1999):

$$R_{rs} = \frac{L_t - \rho L_i}{E_s}, \quad (11)$$

where ρ is the sea surface reflectance factor, which is correlated with the solar zenith angle and surface roughness. Additional details about the R_{rs} calculation can be found in Sokoletsky and Shen (2014). In addition, the ESOA atmospheric correction targets coastal turbid waters. Therefore, we only selected stations with C_{spm} values > 20 (g m^{-3}) (Section 3.1) for comparison and validation. As a result, we selected a total of nine stations, the locations of which are shown in Fig. 5. We optimized four parameters and estimated $R_{rs}(\lambda)$ by applying the ESOA algorithm to the individual pixels of the GOCI images that corresponded to the latitudes and longitudes of the stations. The measurement time difference

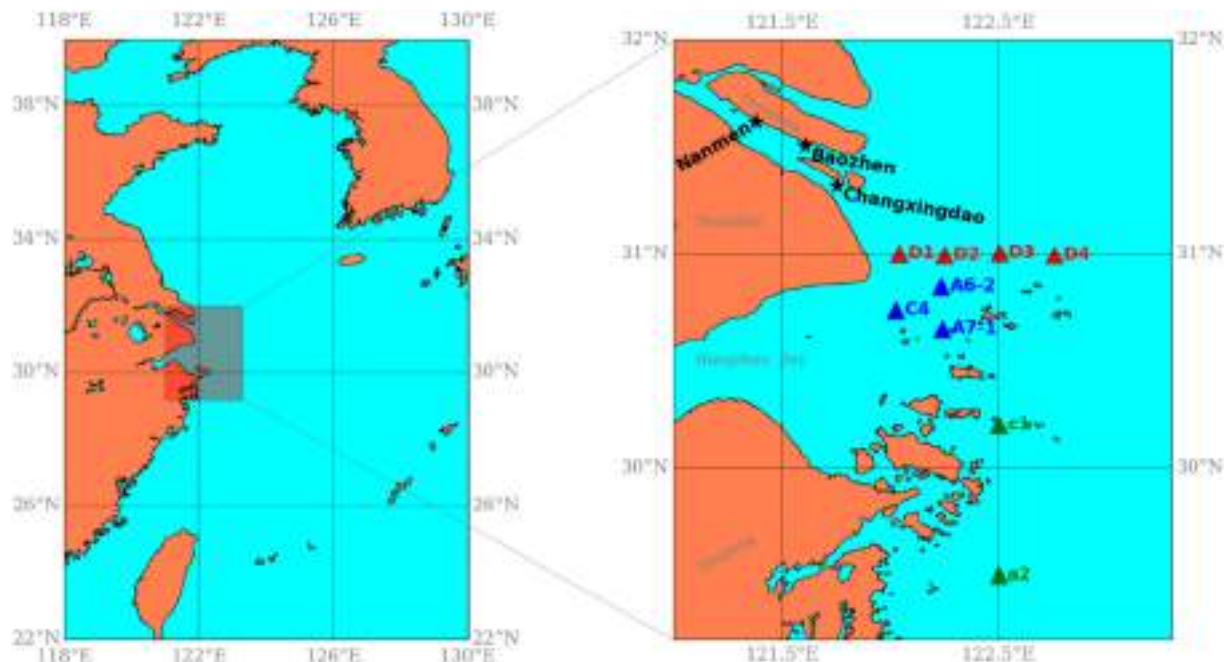


Fig. 5. Locations of matchup datasets collected from cruise stations (triangles) and fixed field stations (stars). Red, purple and green triangles refer to cruises in May 2011, March 2012 and March 2013, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

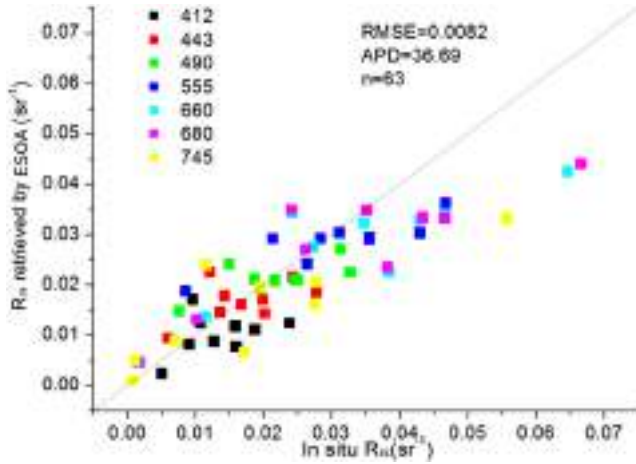


Fig. 6. Comparison between measured and retrieved $R_{rs}(\lambda)$ values for all GOCI wavebands (except for 865 nm). The 1:1 reference line, root mean square error (RMSE), and absolute percentage difference (APD) between the retrieved and measured $R_{rs}(\lambda)$ values are shown.

is <3 h. We evaluated the bias using the RMSE and APD, which are calculated using:

$$RMSE = \sqrt{\frac{\sum_{j=1}^M \sum_{i=1}^N \left[\left(R_{rs}(\lambda_i)^{insitu} - R_{rs}(\lambda_i)^{retrieved} \right)^2 \right]_j}{NM}}, \quad (12)$$

$$APD = 100 \frac{1}{MN} \sum_{j=1}^M \sum_{i=1}^N \left[\frac{|R_{rs}(\lambda_i)^{insitu} - R_{rs}(\lambda_i)^{retrieved}|}{R_{rs}(\lambda_i)^{insitu}} \right]_j, \quad (13)$$

where i and j denote the i th waveband at the j th station, and M and N are the number of stations and the number of bands, respectively. Fig. 6 shows that the estimated and measured $R_{rs}(\lambda)$ values were very consistent and plot close to the 1:1 line. The RMSE and APD were 0.0082 and 36.69, respectively.

4.3. Validation with matchup datasets of SPM

The ESOA algorithm optimizes the four parameters simultaneously. Because it was impossible to find matching measurements for RH, FMF and $\tau_a(865)$, we only performed the comparison validation for C_{spm} . We used the matching datasets from the same cruises as in Section 4.2 to validate the SPM. The SPM concentrations were determined gravimetrically in the laboratory. Water samples were filtered by 0.7 μm Whatman GF/F glass fiber filters. The blank and sample-filled filters were rinsed with Milli-Q water to remove salts, dried, and then reweighed on a high-precision balance in the laboratory (Shen et al., 2014). As shown in Fig. 7, the in situ measured and retrieved SPM concentrations plot close to the 1:1 line. The definitions for the RMSE and APD are similar to those discussed in Section 4.2, and their values were 51.7 and 29, respectively.

In addition to the SPM datasets from the cruises, we collected SPM datasets from fixed field stations where the water turbidity is automatically recorded using D&A Tech optical backscatter instruments (OBSs). Autonomous measurements of OBS turbidity at fixed field stations may increase the opportunity to validate satellite retrieval products. The full-year 2013 datasets from three stations, Changxingdao, Nanmen, and Baozhen, were used for concurrent verification between GOCI and in situ measurements. The locations of the three stations are shown in Fig. 5. GOCI collects one image every other hour between 0 and 8 (UTC) on a daily basis. It takes approximately 30 min for each image scene to scan the Earth's surface. To match the timing of the station data with the GOCI images, we selected the average of the measurements during the scan time of each image as the station measurement. For example, the first GOCI image was acquired from approximately 00:16 (UTC) to 00:45 (UTC). Therefore, we selected the average of the measured station data between 00:16 (UTC) and 00:45 (UTC) as the image matching measurement. Commonly, the OBS directly measures turbidity in units of NTU (defined as x). To obtain the SPM concentration (defined as y in units of $g\ l^{-1}$), we needed to calibrate their relationship. Xue et al. (2004) calibrated the OBS at the Yangtze River estuary and obtained the relationship $y = 0.0017x + 0.0202$. Zhao et al. (2015) performed a calibration for the Chongxi station and found that the relationship between y and x varied between the wet season

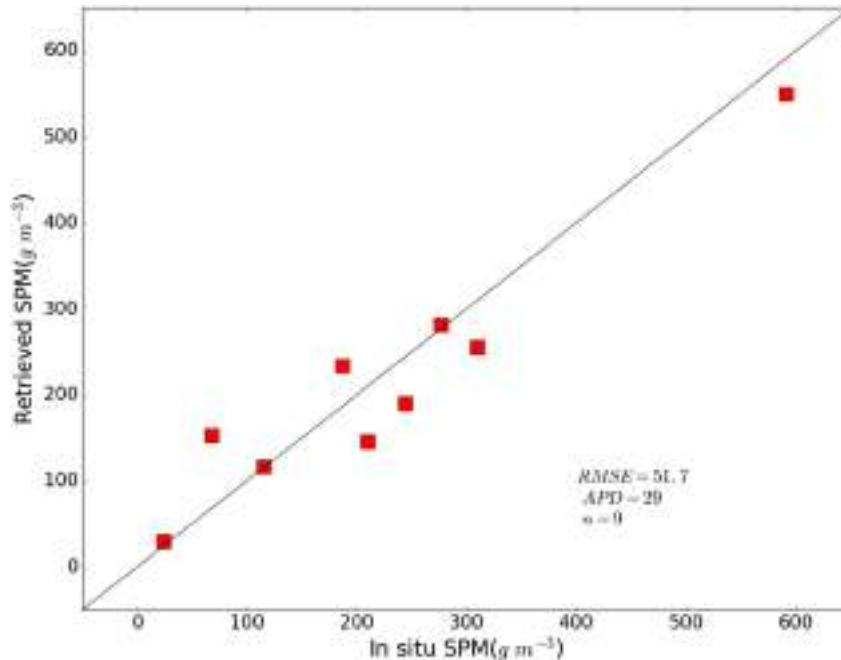


Fig. 7. Comparison between in situ measured and retrieved C_{spm} values ($g\ m^{-3}$). The 1:1 reference line, root mean square error (RMSE), and absolute percentage difference (APD) between the retrieved and in situ measured C_{spm} ($g\ m^{-3}$) values are shown.

(September and October; $y = 0.0021x$) and the dry season (March and April; $y = 0.0013x$). This study used the averages of the SPM concentrations derived from the relationships reported by Xue et al. (2004) and Zhao et al. (2015) as the station SPM concentrations. Using the method presented above, we processed the full-year datasets from the Changxingdao, Nanmen and Baozhen stations that matched the GOCI images. For comparison with the station C_{spm} , we converted the units of C_{spm} from g m^{-3} to g l^{-1} . From the time series of the matching results (Fig. 8), the retrieved C_{spm} values for the three stations reasonably correlated with the SPM variation measured by the OBS. Specifically, for the Changxingdao station, the two values were similar after April 2, but there were large differences for unknown reasons before that. The measured SPM from the Baozhen station was higher than that retrieved from the GOCI data. The largest difference occurred on October 12–19. The GOCI SPM of the Nanmen station deviated considerably from the

measured SPM before April 2. The measured data did not vary after October 29 and September 9 at the Baozhen and Nanmen stations, respectively. This may indicate that the instrument malfunctioned during that time. In general, the OBS turbidity cannot always reliably replace the SPM concentration because the SPM concentration (dry-mass weight per volume) is only proportional to the OBS turbidity (NTU), but the gain factor might vary; it depends on the calibrated relationship between the OBS turbidity and the corresponding SPM concentration, which is not always stable and can be impacted by the particle size, shape and composition, aggregation/flocculation, bubbles and chemical and biological fouling. Unknown temporal variations in the particle size or aggregation/flocculation and fouling will nearly always result in inaccurate OBS data (Downing, 2006). An effective way to develop the calibrated relationship would be to conduct OBS measurements with simultaneous water samples, which may reduce these physical and

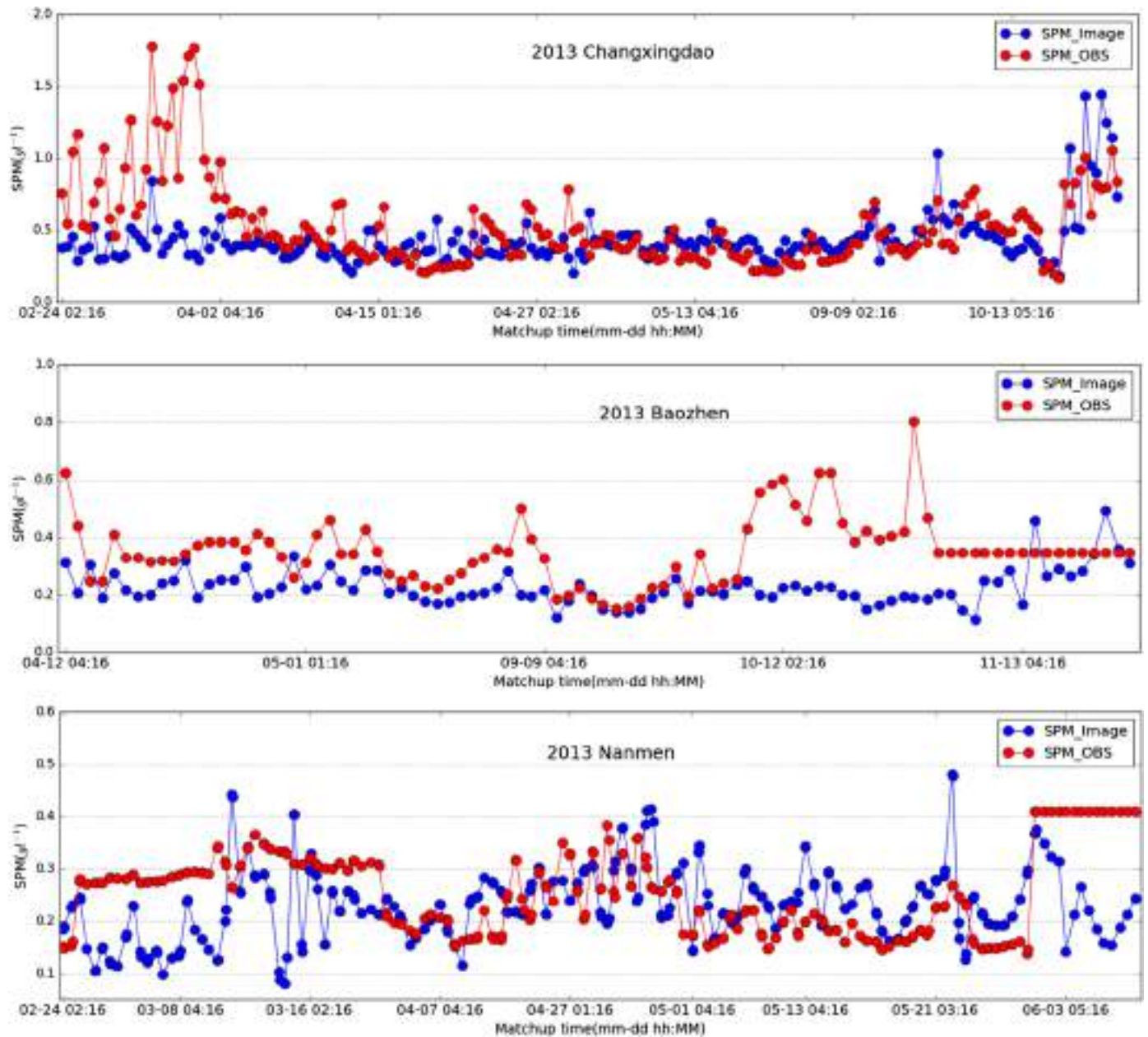


Fig. 8. Synchronous GOCI-derived SPM (blue) obtained using the proposed ESOA algorithm with concurrent OBS SPM at three fixed field stations (Changxingdao, Baozhen and Nanmen) in 2013 (red). There are 202, 90 and 241 matchups for Changxingdao (a), Baozhen (b) and Nanmen (c), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

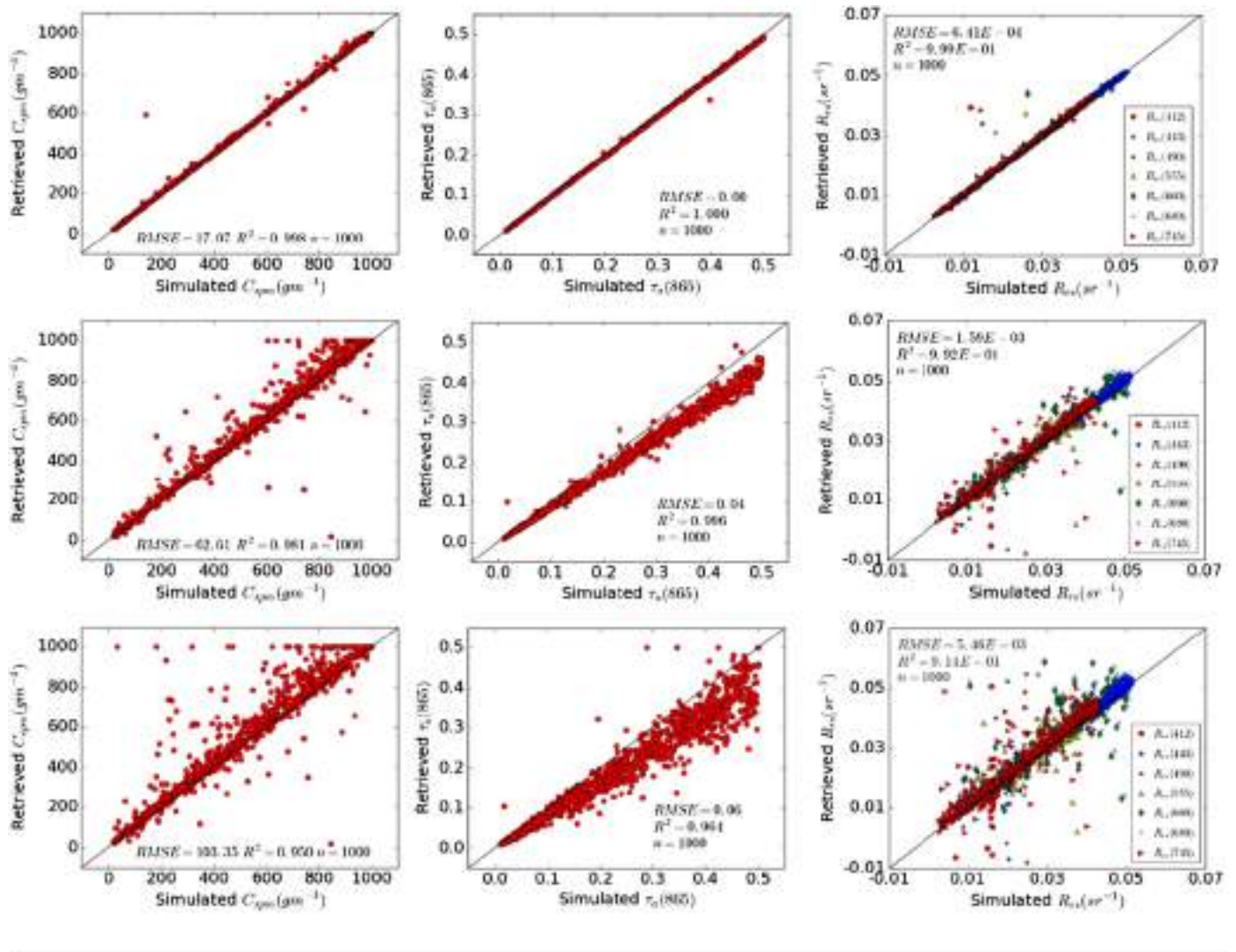


Fig. 9. Retrieved values of SPM, $\tau_a(865)$ and $R_{rs}(\lambda)$ (from left to right) versus the actual values in the synthetic datasets with $RH = 70$, $FMF = 20$ and no deviation (upper panels), 5% deviation (middle panels), and 20% deviation (lower panels).

biochemical impacts. However, this is impractical for autonomous OBS measurements due to their inaccessibility, remote operation and limited personnel. Additional reasons for the deviations may not be completely neglected, such as the difference in footprint between the spatial resolutions of the satellite observations and the in situ measurements and the effect of neighboring land in narrow channels of the estuary.

5. Results and discussion

5.1. The operational satellite image processing approach

The ESOA algorithm uses the Ahmad2010 model as the aerosol model, which can somewhat solve the absorbing-aerosol problem for coastal waters. It uses the SERT model as the water model, which

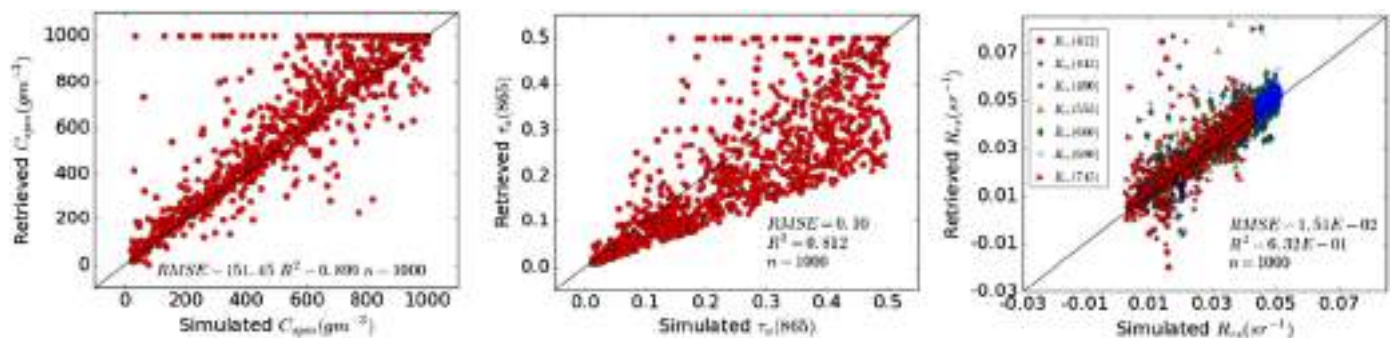


Fig. 10. SPM, $\tau_a(865)$ and $R_{rs}(\lambda)$ values (from left to right) retrieved by LM without the assumption of a homogenous aerosol type or the first estimate from ESOA versus the actual values in the synthetic datasets, which are the same as those in the upper panels in Fig. 9.

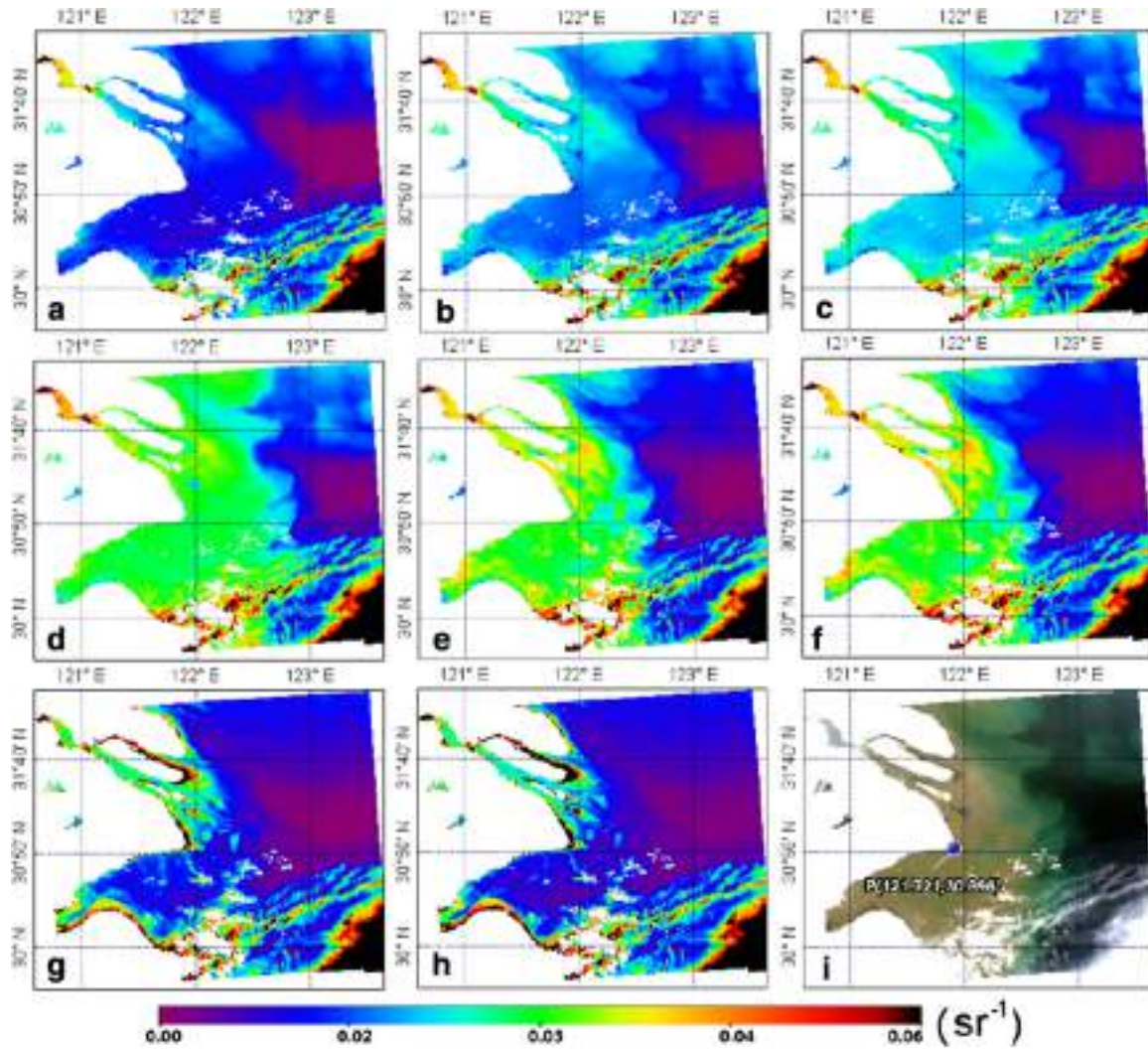


Fig. 11. $R_{rs}(\lambda)$ values retrieved by GOCI using ESOA assuming a spatially homogeneous aerosol type distribution: (a) 412 nm; (b) 443 nm; (c) 490 nm; (d) 555 nm; (e) 660 nm; (f) 680 nm; (g) 745 nm; (h) 865 nm; (i) RGB pseudo color image using TOA reflectance at 660 nm (red), 555 nm (green), and 443 nm (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

simplifies the model for highly turbid waters and facilitates the optimization calculation, and it uses a genetic algorithm to optimize the non-linear problem, which avoids the traditional local optima-searching limitation. The validation with simulated and measured datasets shows that the ESOA algorithm has satisfactory accuracy for atmospheric corrections over turbid waters. However, it takes a long time for ESOA to optimize individual remote-sensing image pixels. The run time for a computer with an Intel i7 processor and 16 GB of memory was approximately 2 s. Thus, approximately 140 h were required to process a 500×500 pixel remote-sensing image, which is unacceptable. A particular technique is used to overcome this problem and apply the ESOA algorithm to operational processing. We assume that the aerosol is spatially homogeneous over the study area, and the satellite imagery is then corrected in two steps:

- 1) Select a cloud-free pixel over the region and optimize the four parameters for that pixel with the ESOA algorithm. The optimization results of the four parameters are denoted RH' , FMF' , C_{spm}' and $\tau_a(865)'$, respectively.
- 2) Assume that the aerosol type is homogenous and that the values of RH and FMF for each pixel are RH' and FMF' , respectively. Thus, each pixel has two remaining unknown parameters (C_{spm} and $\tau_a(865)$). A fast nonlinear optimization algorithm called LM is then

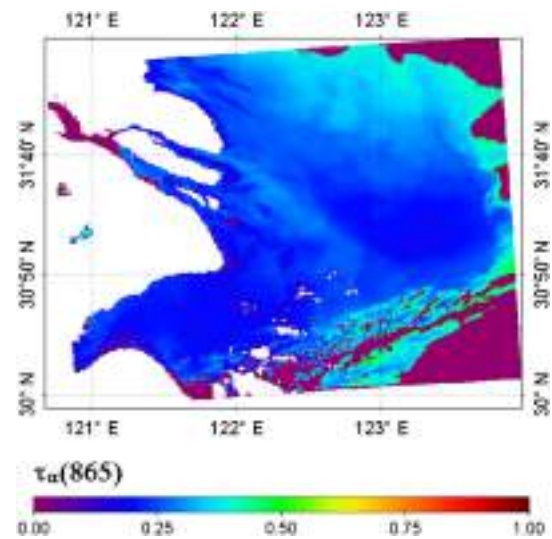


Fig. 12. The $\tau_a(865)$ (m^{-1}) values retrieved by GOCI using ESOA assuming a spatially homogeneous aerosol type distribution. Cloudy pixels are masked in purple. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

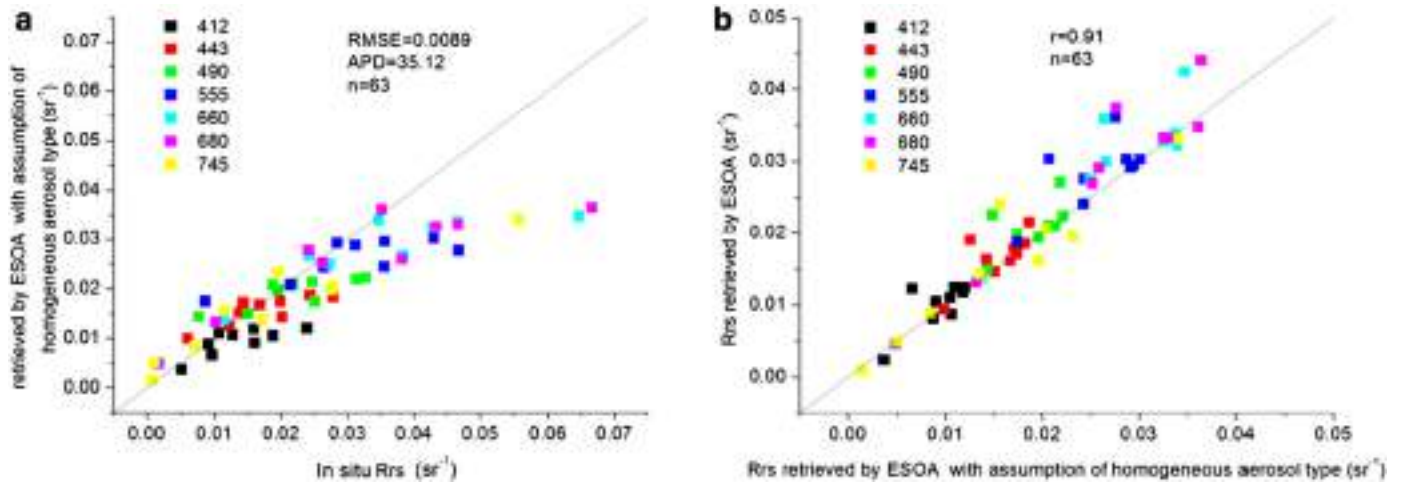


Fig. 13. (a) Comparison between measured and retrieved $R_{rs}(\lambda)$ values of all GOCI wavebands (except for 865 nm) using ESOA with the assumption of a homogeneous aerosol type. The 1:1 reference line, root mean square error (RMSE), and absolute percentage difference (APD) between the retrieved and measured $R_{rs}(\lambda)$ values are shown. (b) Comparison between the $R_{rs}(\lambda)$ values retrieved by ESOA and those retrieved by ESOA with the assumption of a homogeneous aerosol type. The correlation coefficients, r , are shown.

used to optimize C_{spm} and $\tau_a(865)$ with the initial values C_{spm}' and $\tau_a(865)'$, respectively. Finally, $R_{rs}(\lambda)$ is estimated using Eq. (10).

To theoretically evaluate this technique, we simulated a group of datasets with a specific combination of $RH = 70$ and $FMF = 20$ and a randomly generated observation geometry, C_{spm} and $\tau_a(865)$. In addition, to evaluate the retrieval sensitivity to variations of RH and FMF , two additional groups of datasets with different deviations of RH and FMF (5% and 20%) were generated. These three groups of datasets were processed using the technique described above.

For comparison, the synthetic datasets with no deviations of RH and FMF are also optimized by LM directly without the assumption of a homogeneous aerosol type or the first estimate from ESOA but rather with a random initialization. For the synthetic datasets with no deviations of RH and FMF , the retrieved values of C_{spm} , $\tau_a(865)$ and R_{rs} are consistent. As the deviation increases, the retrieval errors of C_{spm} , $\tau_a(865)$ and R_{rs} increase but remain consistent (Fig. 9). The technique has significantly lower retrieval errors than the LM optimization without the assumption of a homogeneous aerosol type or the first estimate from ESOA even for the datasets with 20% deviations of RH and FMF (Fig. 10).

We employed the method described above for the atmospheric correction of one GOCI image (2011.05.07, 02:16 UTC) over the Yangtze estuary and the adjacent coast. The target region includes 540×540 pixels (including the continent), for which the processing time was 15 min. Fig. 11(a)–(h) show images at each waveband after the correction. The aerosol type of point P in Fig. 11(i) was used as the fixed aerosol type for the entire target region. Fig. 12 shows the retrieved $\tau_a(865)$ image.

To further validate this technique, we processed GOCI images over the Yangtze estuary and the adjacent coast from the matchup datasets in Sections 4.2 and 4.3 and estimated the $R_{rs}(\lambda)$ values. The comparison with the measurements showed that the derived $R_{rs}(\lambda)$ values were quite accurate with the assumption of a homogeneous aerosol. Although the RMSE (~ 0.0089) was slightly higher than the RMSE without this assumption (~ 0.0082 ; Fig. 6), the APD (35.12) was slightly lower. Overall, there was no increase in the error (Fig. 13(a)). Fig. 13(b) shows that the retrieved $R_{rs}(\lambda)$ values with and without the assumption of a homogeneous aerosol were well correlated with a correlation coefficient r of up to 0.91. These results are generally consistent with those from the simulated datasets. Therefore, we conclude that the assumption of a homogeneous aerosol type over the Yangtze estuary and the adjacent coast does not significantly affect the accuracy of the $R_{rs}(\lambda)$ retrieval.

5.2. Comparison with GDPS

In addition, to compare the atmospheric correction algorithm to the official standard atmospheric correction algorithm of GOCI imagery, we employed GDPS (v1.3, http://kosc.kiost.ac/kosc_web/?sub_num=23) for the atmospheric correction of the GOCI images from the matchups in Sections 4.2 and 4.3. The comparison between the GDPS v1.3-corrected and measured $R_{rs}(\lambda)$ values showed that the atmospheric correction algorithm that is based on ESOA and the assumption of a homogeneous aerosol was more accurate (Fig. 14) over the Yangtze estuary and the adjacent coast. The RMSE and APD for the latter were 0.0089 and 35.12, respectively (Fig. 13(a)), whereas they were 0.0104 and 69.15 for the former, which are 17% and 97% greater, respectively.

5.3. Application to MERIS

The ESOA algorithm is applicable not only to GOCI images but also to other ocean color sensors, such as MERIS. We applied the ESOA algorithm with the assumption of a homogeneous aerosol to the 2011.5.7

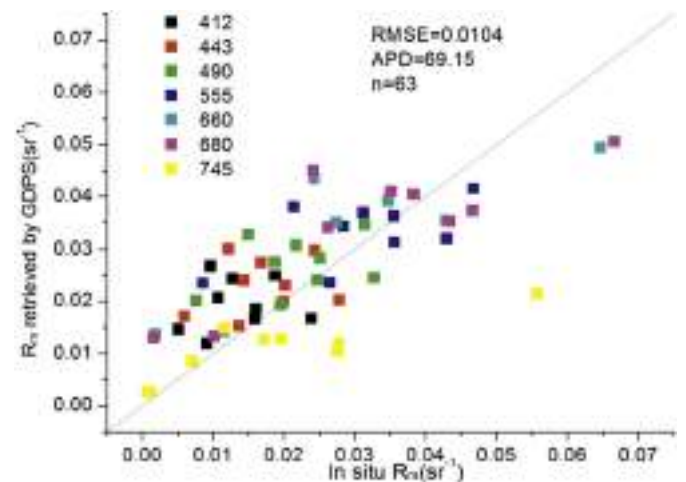


Fig. 14. Comparison between measured and retrieved $R_{rs}(\lambda)$ values of all GOCI wavebands (except for 865 nm) using GDPS (v1.3). The 1:1 reference line, root mean square error (RMSE), and absolute percentage difference (APD) between the retrieved and measured R_{rs} values are shown.

02UTC MERIS image over the Yangtze estuary and the adjacent coast. The neural-network (NN) atmospheric correction algorithm (Schroeder et al., 2007) in BEAM (v4.11, <http://www.brockmann-consult.de/cms/web/beam/releases>) was also used to process the image. Fig. 15 shows the MERIS-derived $R_{rs}(\lambda)$ images at a selection of MERIS wavebands using the two atmospheric correction algorithms. In the Yangtze estuary and the adjacent coast, the corrected $R_{rs}(\lambda)$ values using the NN algorithm were generally lower than those using ESOA at all MERIS wavebands; this was most significant at 665 nm and 681 nm (Fig. 15). We compared the corrected $R_{rs}(\lambda)$ images at the D3 station (location shown in Fig. 5) with the two algorithms and with the in situ $R_{rs}(\lambda)$ measurements. As shown in Fig. 16, the $R_{rs}(\lambda)$ values from the two algorithms were both underestimated at visible wavelengths, but the values from the ESOA algorithm were closer to the measured values. At 665 nm and 681 nm, the ESOA algorithm performed much better than the NN algorithm. Of course, the training samples used in the NN algorithm may have come from European coastal regions, which may have caused the lower performance in the extremely turbid waters of the Yangtze estuary and the adjacent coast. The purpose of the comparison is to show the performance of ESOA when it is applied to MERIS images. Although ESOA with the assumption of a

homogeneous aerosol performed well in the $R_{rs}(\lambda)$ retrieval over the Yangtze estuary and the adjacent coast, this algorithm is more applicable to small-scale regions. For large-scale regions, larger errors may occur when using a uniform aerosol type to represent different aerosol types over an entire region due to possible variations in the aerosol types.

5.4. Sensitivity analysis

The potentially weak points of this method are related to the sources of uncertainty in the approach and possible gaps that might still be present. Therefore, we performed an additional sensitivity analysis, and the results are presented in this section. To analyze the sensitivity of the reflectance spectra to variations in Chla concentration and CDOM, we simulated R_{rs} under two scenarios. In the first scenario, we assume that $a_{cdm}(440)$ is 5 m^{-1} , the SPM concentration varies over 20, 200 and 500 g m^{-3} , and Chla is $\sim[1, 3, 5, 7, 9] \text{ mg m}^{-3}$. The R_{rs} spectra can be simulated by the two-stream Kubelka-Munk model (Eq. (4)) and several other models, such as that in Lee et al. (1998). In the second scenario, we assume that Chla is 5 mg m^{-3} , the SPM concentration varies over 20, 200 and 500 g m^{-3} , and $a_{cdm}(440)$ is $\sim[0.1, 0.3, 0.5, 0.7, 0.9] \text{ m}^{-1}$,

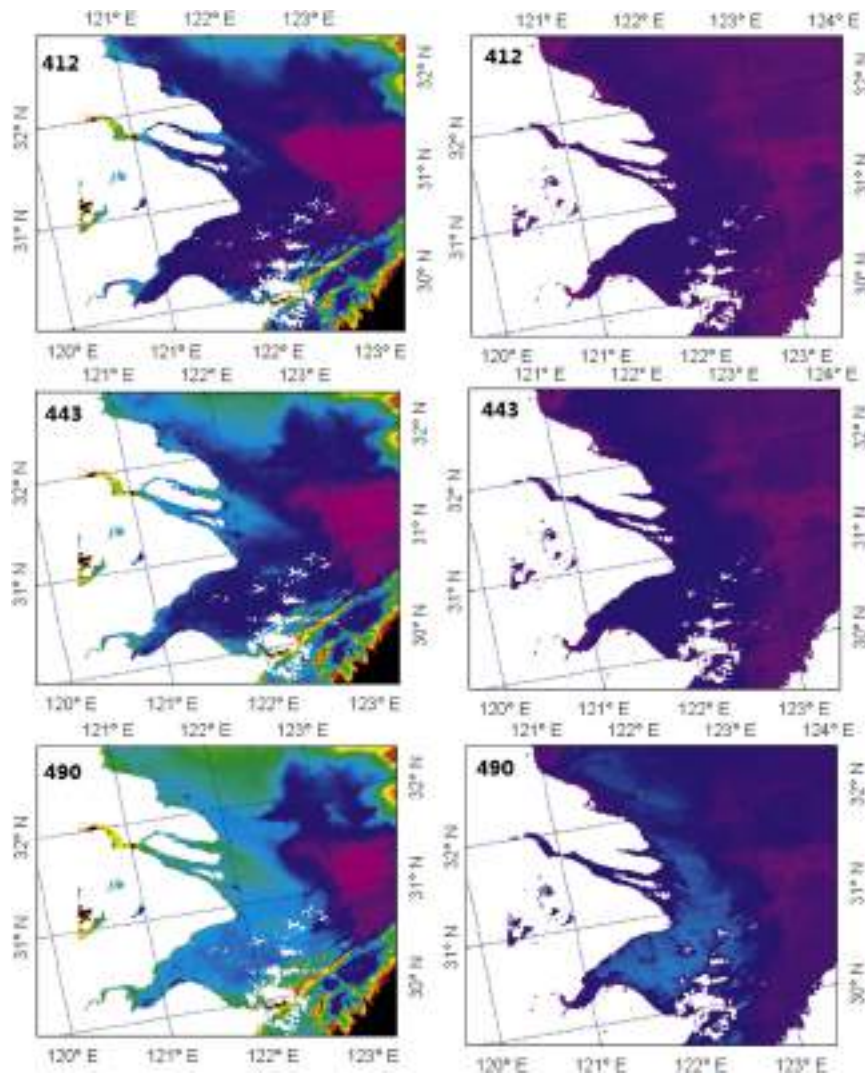


Fig. 15. Comparison of MERIS $R_{rs}(\lambda)$ images at wavebands of 412, 443, 490, 560, 665, 754, and 779 nm on May 7, 2011. The results in the left panels are derived from the atmospheric correction algorithm that is based on ESOA, and those in the right panels are derived from the atmospheric correction algorithm that is based on the NN. Station D3 is also shown (for 560 nm).

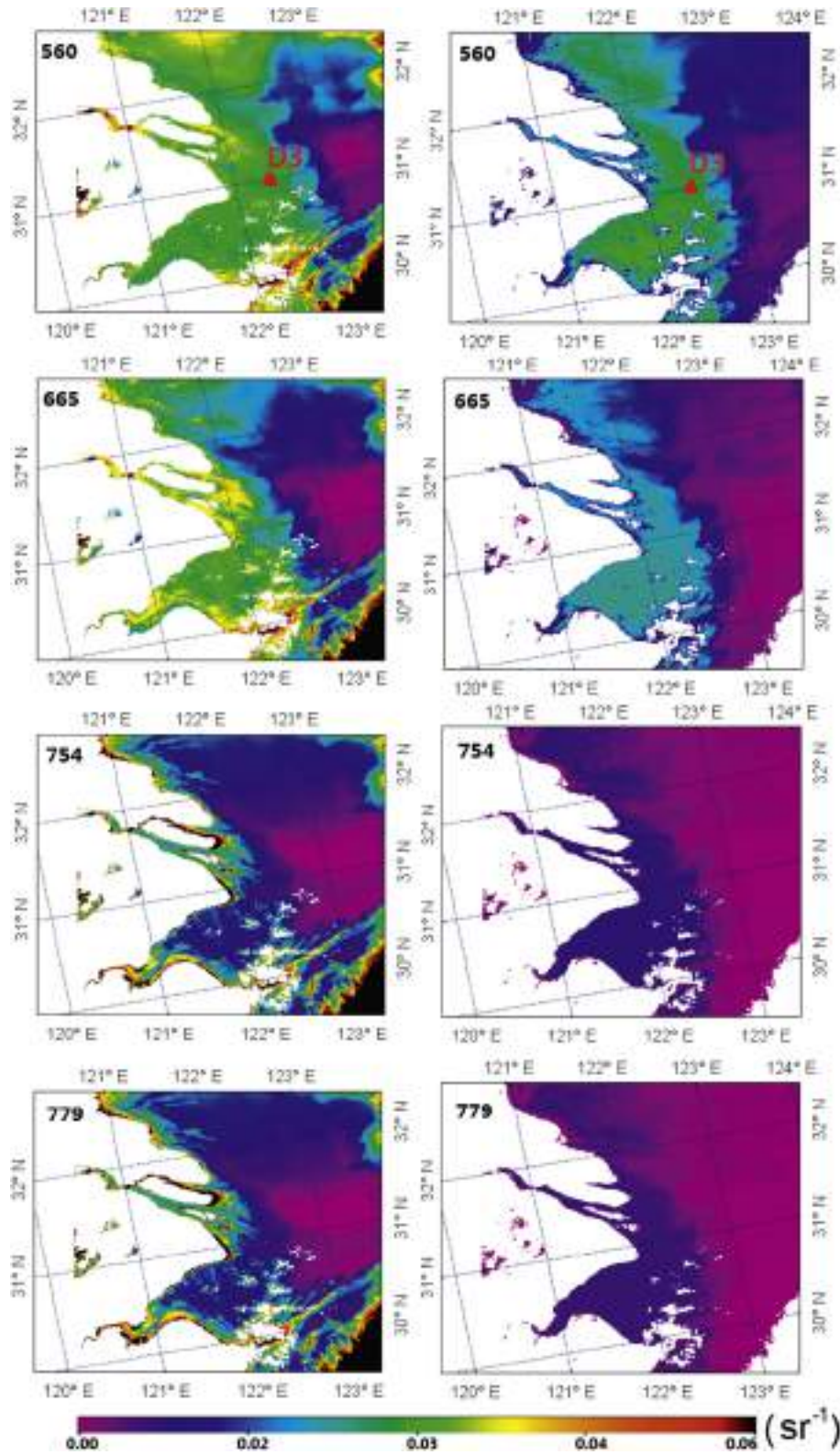


Fig. 15 (continued).

and the R_{rs} spectra were simulated. The relative change rate (RCR) of R_{rs} , which indicates the sensitivity to the variations of Chla (scenario 1) and CDOM (scenario 2), was estimated for the two scenarios by:

$$RCR = \frac{R_{rs}^i - R_{rs}^0}{R_{rs}^0} 100\%, \quad (14)$$

where R_{rs}^i refers to the R_{rs} spectra with different values of Chla and CDOM, and R_{rs}^0 refers to the base R_{rs} with fixed Chla and CDOM. For the two scenarios, R_{rs} with $a_{cdom}(440)$ of 0.5 m^{-1} and Chla of 5 mg m^{-3} is selected as the base R_{rs}^0 .

Figs. 17 and 18 show that the high RCR values mainly occur at shorter wavelengths (e.g., the blue bands). When the SPM concentration increases, the RCR decreases. This means that the R_{rs} spectra may

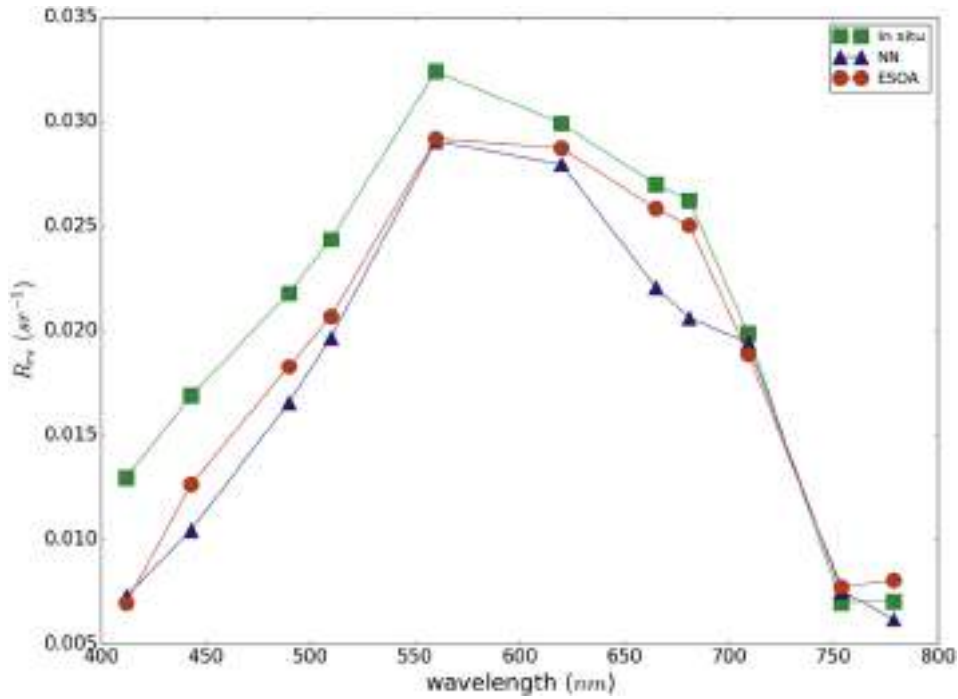


Fig. 16. $R_{rs}(\lambda)$ spectra derived from the two atmospheric correction algorithms based on ESOA and the NN for station D3 (location shown in Fig. 5) for a MERIS image acquired on May 7, 2011. The in situ $R_{rs}(\lambda)$ spectrum is also shown for comparison.

not be sensitive to variations of the Chla and CDOM components when the waters with high SPM become turbid, especially at longer wavelengths. Therefore, the SERT model, which is based on the assumption that the concentrations of Chla and CDOM are constant, is a reasonable approximation for highly turbid waters.

To analyze the impacts of Chla and CDOM variations on the retrievals of SPM, $\tau_a(865)$ and R_{rs} , we simulated ρ_m spectra at the GOCI bands by coupling these simulated R_{rs} datasets with the Ahamd2000 aerosol model. The RH, FMF and $\tau_a(865)$ values in the aerosol model are 70, 10 and 0.1, respectively, and the angular geometry is 30, 30 and 90. The optimization and inversion were processed by the ESOA. The relative error (RE) of the inversion is estimated by:

$$RE = \frac{P_{\text{retrieval}} - P_{\text{true}}}{P_{\text{true}}} 100\%, \quad (15)$$

where $P_{\text{retrieval}}$ is the inverted value, and P_{true} is the true value.

Fig. 19 shows the impacts of variations in the Chla concentration and CDOM on the retrievals of C_{spm} , $\tau_a(865)$ and R_{rs} at the GOCI wavebands. At

a constant SPM concentration, the smaller the change in Chla is, the smaller the inversion error (Fig. 19(a)). If Chla is less than the assumed value of 5, SPM is underestimated; otherwise, SPM is overestimated. The error of the SPM inversion decreases with increasing SPM concentration. Compared to Chla, the same percentage variation in CDOM has a greater influence on the SPM retrieval. For $\tau_a(865)$ (Fig. 19(b)), the inversion error does not show a clear pattern; when SPM is 20 g m^{-3} , the $\tau_a(865)$ values at different Chla concentrations are underestimated. The Chla variations cause less than a 20% RE in the retrieval of R_{rs} (Fig. 19(c)–(i)). The CDOM variations cause a greater RE in its retrieval than Chla. The R_{rs} inversion error is also related to the variation in CDOM; the smaller the change in CDOM is, the smaller the inversion error. Similar to SPM, the inversion RE of R_{rs} decreases with increasing SPM concentration. In addition, the influences on short wavelengths (412 and 443 nm) are greater; the largest RE of R_{rs} at 412 nm (Fig. 19(c)) is $>50\%$, whereas the inversion REs of R_{rs} at the other wavebands are generally $<20\%$.

The results presented in this section confirm that the semi-empirical SERT model can be successfully applied over waters with high sediment concentrations, although varying concentrations of Chla and CDOM

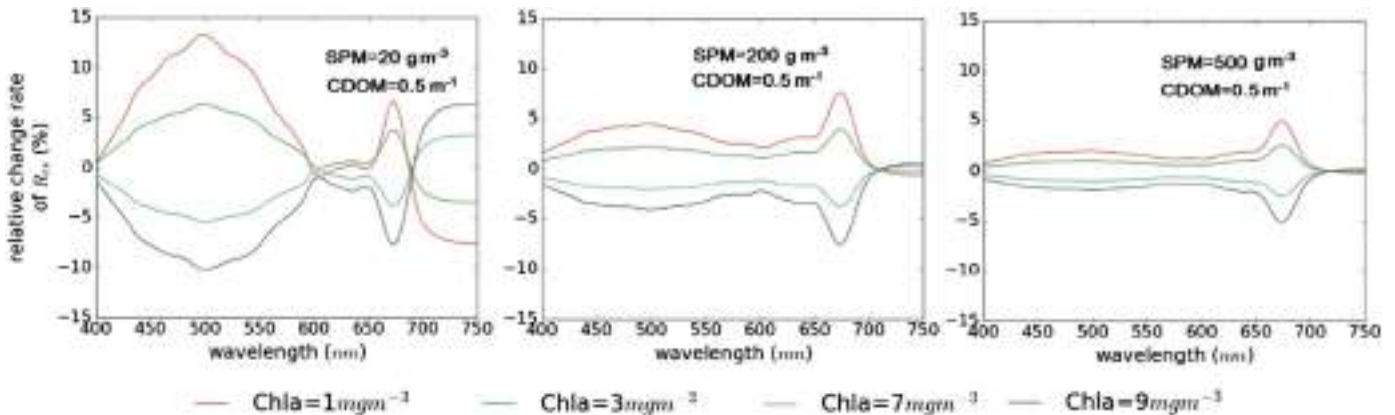


Fig. 17. The impact of changes in Chla concentration on R_{rs} . (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

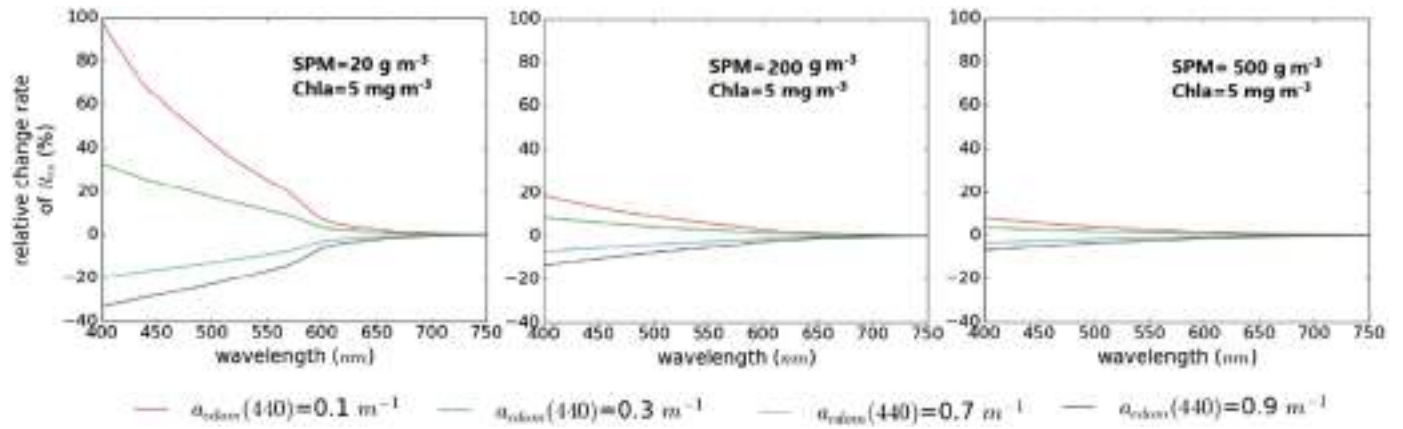


Fig. 18. The impact of changes in CDOM on R_{rs} . (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

might affect the linear relationship between the factor $x = b_b/a$ and the sediment concentration, especially in the blue-green parts of the spectrum. In addition, when coupling the SERT model with the aerosol model for atmospheric correction and parameter inversion, the results demonstrate that although variations in Chla and CDOM have significant influences on the $\tau_a(865)$ retrieval ($|RE| > 100\%$), they do not have a significant influence on the C_{spm} ($|RE| < 25\%$) and R_{rs} ($|RE| < 60\%$) retrievals.

6. Conclusion

This study proposed the ESOA algorithm based on several limitations of the SOA algorithm for atmospheric corrections over turbid

coastal waters. The ESOA algorithm combines Ahmad2010 as the aerosol model and SERT as the water model and applies a genetic algorithm for the nonlinear equation optimization. Based on the validation of a simulated dataset and matchups between GOCI images and in situ measurements (including fixed station measurements), the algorithm generated good optimizations for C_{spm} and $\tau_a(865)$ and was more accurate for the $R_{rs}(\lambda)$ retrieval. The major disadvantage of the ESOA algorithm is that it is very time-consuming. To overcome this shortcoming and apply this algorithm to processing satellite images, a technique that is based on ESOA, assumes a spatially homogenous aerosol type and uses the fast optimization algorithm LM is proposed. Although the aerosol type over coastal regions can vary spatially, the validation results show that this had little effect on the $R_{rs}(\lambda)$

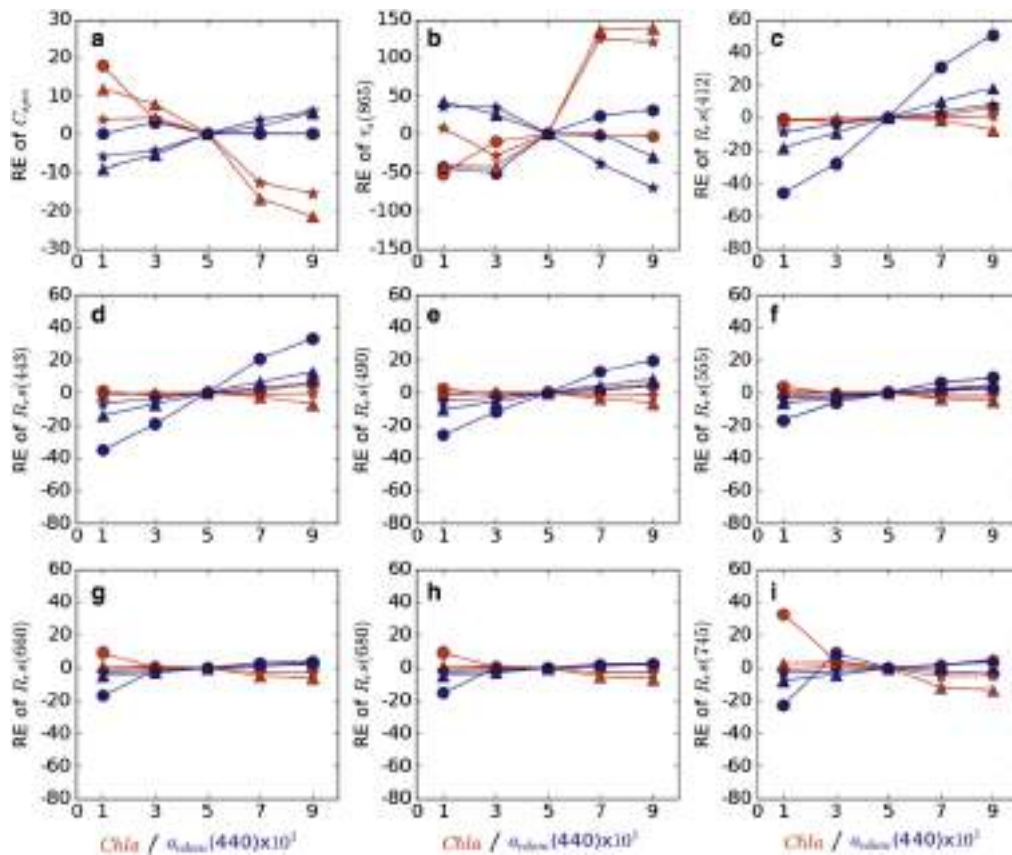


Fig. 19. Relative errors (RE in %) of the influences of Chla (red) and CDOM (blue) on the retrieved values of C_{spm} (a), $\tau_a(865)$ (b) and R_{rs} at GOCI wavebands: (c) 412 nm; (d) 443 nm; (e) 490 nm; (f) 555 nm; (g) 660 nm; (h) 680 nm; (i) 745 nm. Chla and CDOM are shown on the same X-axis but have different units (mg m^{-3} and m^{-1} , respectively) and different scales (1 and 10, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

inversion over the Yangtze River estuary and the adjacent coast. The ESOA algorithm with this technique is more accurate for $R_{rs}(\lambda)$ retrieval than the official GOCI standard atmospheric correction algorithm GDPS. ESOA with this technique can also be applied to other satellite images, such as MERIS. Thus, ESOA is an ideal atmospheric correction method for satellite images over the Yangtze estuary and the adjacent coast.

Acknowledgments

This work was supported by the National Science Foundation of China (no. 41271375 and no. 41576092). The later stages were funded by the 111 Project and Oversea Foundation of East China Normal University and the State Key Laboratory of Estuary and Coastal Research. The authors would like thank KORDI/KOSC for providing the GOCI data, the European Space Agency (ESA) for providing the MERIS image, and all of the participants for collecting the Yangtze datasets. We thank NASA for the SeaDAS processing software and its open source code. Some of the image analyses were performed with the open-source NanSat Python package (<https://github.com/nanscenter/nansat>).

Appendix A. Implementation details of the GA

A.1. The ranges and encoding schemes of variables

The nonlinear problem in this study includes four unknown variables: the relative humidity (RH), fine-mode fraction (FMF), aerosol optical thickness $\tau_a(\lambda_0)$ at the reference wavelength λ_0 and SPM concentration (C_{spm}). The generic algorithm first needs to compose the variables into individual sequences via an encoding scheme. Common schemes include binary encoding and real encoding. Real encoding has the advantage of easy handling, and all of the variables in this study's nonlinear problem are real numbers. Therefore, the individual encoding uses real encoding. The ranges of RH and FMF are 30%–95% and 0–100, respectively. The aerosol optical thickness varies significantly with wavelength. Therefore, we selected the longer near-infrared wavelengths of the target sensors as the reference wavelengths λ_0 . The reference wavelengths λ_0 for MERIS and GOCI are 869 and 865 nm, respectively. According to the AERONET observation data, the aerosol optical thickness at 865 nm is approximately 0.01–0.5. For turbid coastal waters, the SPM concentrations vary over a wide range. Shen et al. (2010a,b, 2013) classified water turbidity according to SPM concentration; specifically, the low, medium, high and ultra-high turbidities are $C_{spm} < 20$ (g m^{-3}), 20 (g m^{-3}) $< C_{spm} < 80$ (g m^{-3}), 80 (g m^{-3}) $< C_{spm} < 250$ (g m^{-3}) and > 250 (g m^{-3}), respectively. The atmospheric correction over low turbidity water (clear water) is beyond the scope of this study. Therefore, this study set the range of C_{spm} to 20–1000 (g m^{-3}). The ranges of the four variables are shown in Table A1.

Table A1
Ranges of the four variables.

Variable	Minimum	Maximum	Unit
RH	30	95	%
FMF	0	100	
$\tau_a(\lambda_0)$	0.01	0.5	
C_{spm}	20	1000	g m^{-3}

A.2. Objective function and fitness

The goal of spectral optimization is to match the $\rho_m(\lambda)$ spectrum to the $\rho_w(\lambda, \text{RH}, \text{FMF}, \tau_a(\lambda_0), C_{spm})$ spectrum as closely as possible. Therefore, the RMSE is usually selected as the objective function (Eq. (A1)). In addition to the RMSE, the correlation coefficient r of the $\rho_m(\lambda)$ spectrum

and the $\rho_w(\lambda, \text{RH}, \text{FMF}, \tau_a(\lambda_0), C_{spm})$ spectrum is an additional objective function (Eq. (A2)).

$$\begin{aligned} p_i &= \rho_m(\lambda_i), \\ q_i &= \rho_{aw}(\lambda_i, \text{RH}, \text{FMF}, \tau_a(\lambda_0), C_{spm}), \\ \text{obj1} &= \sqrt{\frac{\sum_{i=1}^N (p_i - q_i)^2}{N}}, \end{aligned} \quad (\text{A1})$$

$$\text{obj2} = \frac{N \sum_{i=1}^N p_i q_i - \sum_{i=1}^N p_i \sum_{i=1}^N q_i}{\sqrt{N \sum_{i=1}^N p_i^2 - \left[\sum_{i=1}^N p_i \right]^2} \sqrt{N \sum_{i=1}^N q_i^2 - \left[\sum_{i=1}^N q_i \right]^2}}, \quad (\text{A2})$$

where obj1 and obj2 denote objective function 1 and objective function 2, respectively. As stated above, the two objective functions are the RMSE (Eq. (A1)) and the correlation coefficient r (Eq. (A2)). The objective of the former is the minimum value (0), and that of the latter is the maximum value (1). For the objective function with a minimum goal, the fitness is its negative $fit = -1 \times \text{obj1}$, where fit refers to the fitness.

A.3. Genetic operators and parameters

Generic algorithms generally include three types of operators, including selection, crossover and mutation, and several important parameters, such as population size (P), crossover rate and mutation rate. Several genetic operation schemes are available, such as roulette wheel (Miller and Goldberg, 1995), tournament selection (Miller and Goldberg, 1995), one- or two-point and uniform crossover and Gaussian mutation. However, the selection of genetic operators is usually problem specific, and different combinations of genetic operators directly affect the performance of the GA. The issue of genetic operators is not the focus of this study and is not discussed here. This paper uses the genetic operators that were proposed in NSGA-II (Deb et al., 2002), which has been demonstrated to find a much better range of solutions and better convergence for most problems. The GA for the specific problem in this paper is implemented with DEAP (Gagn, 2012), which is a novel evolutionary computational framework for rapid prototyping and testing of ideas (<https://github.com/deap>). Table A2 shows the details of the algorithm implementation, which includes the genetic operation and several key control parameters. To increase the probability of finding the globally optimal solution, a larger population size and the maximum number of generations are applied. The selection, crossover and mutation operators in DEAP are “selNSGA2”, “cxSimulatedBinaryBounded” and “muPolynomialBounded”, respectively. These three genetic operators are implemented based on NSGA-II. The parameter “eta” is the parameter in the crossover and mutation operator; a high “eta” value will produce children that resemble their parents.

Table A2
Details of the genetic operators and parameters in this study.

Parameter	Value
Population size	300
Maximum number of generations	150
Probability of crossover	0.8
Chromosome coding	Real-number coding
Selection operator	selNSGA2
Crossover operator	cxSimulatedBinaryBounded
eta = 20	
Mutation operator	muPolynomialBounded
eta = 20	

References

- Ahmad, Z., Franz, B.A., McClain, C.R., Kwiatkowska, E.J., Werdell, J., Shettle, E.P., Holben, B.N., 2010. 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.* 49 (29), 5545–5560.

- Bailey, S.W., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Opt. Express* 18 (7), 7521–7527.
- Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* 16 (5), 1190–1208.
- Chomko, R.M., Gordon, H.R., 1998. Atmospheric correction of ocean color imagery: use of the Junge power-law aerosol size distribution with variable refractive index to handle aerosol absorption. *Appl. Opt.* 37 (24), 5560.
- Chomko, R.M., Gordon, H.R., Maritorena, S., Siegel, D.A., 2003. Simultaneous retrieval of oceanic and atmospheric parameters for ocean color imagery by spectral optimization: a validation. *Remote Sens. Environ.* 84 (102C), 208–220.
- Davies, C.N., 1974. Size distribution of atmospheric particles. *J. Aerosol Sci.* 5 (3), 293–300.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197.
- Downing, John, 2006. Twenty-five years with OBS sensors: the good, the bad, and the ugly. *Cont. Shelf Res.* 26, 2299–2318.
- Gagn, C., 2012. DEAP: evolutionary algorithms made easy. *J. Mach. Learn. Res.* 13, 2171–2175.
- Garver, S.A., Siegel, D.A., 1997. Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation I. *J. Geophys. Res.* 102 (C8), 18607–18625.
- Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Appl. Opt.* 33 (3), 443–452.
- Gordon, H.R., Brown, J.W., Evans, R.H., 1988. Nimbus-7 coastal zone color scanner. *Appl. Opt.* 27 (5), 862–871.
- Gordon, H.R., Du, T., Zhang, T., 1997. Remote sensing of ocean color and aerosol properties: resolving the issue of aerosol absorption. *Appl. Opt.* 36 (33), 8670–8684.
- He, X., Bai, Y., Pan, D., Tang, J., Wang, D., 2012. Atmospheric correction of satellite ocean color imagery using the ultraviolet wavelength for highly turbid waters. *20* (18), 741–753.
- Hooker, S.B., Lazin, G., Zibordi, G., Mclean, S., 2002. An evaluation of above- and in-water methods for determining water-leaving radiances. *J. Atmos. Ocean. Technol.* 19 (4): 486–515. [http://dx.doi.org/10.1175/1520-0426\(2002\)](http://dx.doi.org/10.1175/1520-0426(2002)).
- Kanzow, C., Yamashita, N., Fukushima, M., 2005. Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *J. Comput. Appl. Math.* 173 (2):321–343. <http://dx.doi.org/10.1016/j.cam.2004.03.015>.
- Knaeps, E., Dogliotti, A.I., Raymaekers, D., Ruddick, K., Sterckx, S., 2012. In situ evidence of non-zero reflectance in the OLCI 1020 nm band for a turbid estuary. *Remote Sens. Environ.* 120, 133–144.
- Kubelka, P., Munk, F., 1931. Ein beitrag zur optik der farbanstriche. *Zeritschrift fr Technische Physik* 12, 593–601.
- Kuchinke, C.P., Gordon, H.R., Franz, B.A., 2009. Spectral optimization for constituent retrieval in Case 2 waters I: implementation and performance. *Remote Sens. Environ.* 113 (3), 571–587.
- Lee, Z., Carder, K.L., Mobley, C.D., et al., 1998. Hyperspectral remote sensing for shallow waters. I. A semianalytical model. *Appl. Opt.* 37 (27), 6329–6338.
- Mao, Z., Chen, J., Hao, Z., Pan, D., Tao, B., Zhu, Q., 2013. A new approach to estimate the aerosol scattering ratios for the atmospheric correction of satellite remote sensing data in coastal regions. *Remote Sens. Environ.* 132, 186–194.
- Mao, Z., Pan, D., Hao, Z., Chen, J., Tao, B., Zhu, Q., 2014. A potentially universal algorithm for estimating aerosol scattering reflectance from satellite remote sensing data. *Remote Sens. Environ.* 142, 131–140.
- Maritorena, S., Siegel, D.A., Peterson, A.R., 2002. Optimization of a semi-analytical ocean color model for global-scale applications. *Appl. Opt.* 41 (15), 2705–2714.
- Miller, B.L., Goldberg, D.E., 1995. Genetic algorithms, tournament selection, and the effects of noise. *Complex Syst.* 9 (3), 193–212.
- Mobley, C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Opt.* 38 (36):7442–7455. <http://dx.doi.org/10.1364/AO.38.007442>.
- Mobley, C.D., Werdell, J., Franz, B., Ahmad, Z., Bailey, S., 2016. Atmospheric Correction for Satellite Ocean Color Radiometry. A Tutorial and Documentation NASA Ocean Biology Processing Group. pp. 1–73.
- Ruddick, K.G., Ovidio, F., Rijkeboer, M., 2000. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl. Opt.* 39 (6), 897–912.
- Schroeder, T., Behnert, I., Schaale, M., Fischer, J., Doerffer, R., 2007. Atmospheric correction algorithm for MERIS above case-2 waters. *Int. J. Remote Sens.* 28 (7), 1469–1486.
- Shen, F., SuhybSalama, M., Zhou, Y.-X., Li, J.-F., Su, Z., Kuang, D.B., 2010a. Remote-sensing reflectance characteristics of highly turbid estuarine waters – a comparative experiment of the Yangtze River and the Yellow River. *Int. J. Remote Sens.* 31 (10), 2639–2654.
- Shen, F., Verhoef, W., Zhou, Y., Salama, M.S., Liu, X., 2010b. Satellite estimates of wide-range suspended sediment concentrations in Changjiang (Yangtze) estuary using MERIS data. *Estuar. Coasts* 33 (6), 1420–1429.
- Shen, F., Zhou, Y., Li, J., He, Q., Verhoef, W., 2013. Remotely sensed variability of the suspended sediment concentration and its response to decreased river discharge in the Yangtze estuary and adjacent coast. *Cont. Shelf Res.* 69, 52–61.
- Shen, F., Zhou, Y., Peng, X., Chen, Y., 2014. Satellite multi-sensor mapping of suspended particulate matter in turbid estuarine and coastal ocean, China. *Int. J. Remote Sens.* 35 (11–12), 4173–4192.
- Shettle, E.P., Fenn, R.W., 1979. Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variations on Their Optical Properties. Air Force Laboratory, Hanscom Air Force Base, Mass, AFGL-TR 790214, U. S.
- Shi, W., Jiang, L., Wang, M., 2012. Atmospheric correction using near-infrared bands for satellite ocean color data processing in the turbid western Pacific region. *Opt. Express* 20 (2), 741.
- Siegel, D.A., Wang, M., Maritorena, S., Robinson, W., 2000. Atmospheric correction of satellite ocean color imagery: the black pixel assumption. *Appl. Opt.* 39 (21), 3582–3591.
- Sokoletsky, L.G., Shen, F., 2014. Optical closure for remote-sensing reflectance based on accurate radiative transfer approximations: the case of the Changjiang (Yangtze) River Estuary and its adjacent coastal area, China. *Int. J. Remote Sens.* 35 (11–12): 4193–4224. <http://dx.doi.org/10.1080/01431161.2014.916048>.
- Stramska, M., Petelski, T., 2003. Observations of oceanic whitecaps in the north polar waters of the Atlantic. *J. Geophys. Res.* 108 (C3):3086–3095. <http://dx.doi.org/10.1029/2002JC001321>.
- Wang, M., Bailey, S.W., 2001. Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products. *Appl. Opt.* 40 (27):4790. <http://dx.doi.org/10.1364/AO.40.004790>.
- Wang, M., Shi, W., 2005. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing Menghua. *J. Geophys. Res.* 110 (D10), D10S06.
- Wang, M., Shi, W., Son, S., 2009. Evaluation of MODIS SWIR and NIR-SWIR atmospheric correction algorithms using SeaBASS data. *Remote Sens. Environ.* 113 (3), 635–644.
- Xue, Y., He, Q., Wang, Y., 2004. The method and application of OBS in the measurement of sediment concentration. *Chin. J. Sediment. Res.* 4, 56–60.
- Zhao, J., He, Q., Wang, X., Guo, L., 2015. Field observations on the characteristics of current and sediment of the south and north branches in the Yangtze estuary, Chinese. *Resources and Environment in Yangtze Basin* 24 (1), 21–27.