Semianalytical Derivation of Phytoplankton, CDOM, and Detritus Absorption Coefficients From the Landsat 8/OLI Reflectance in Coastal Waters

Jianwei Wei1, ZhongPing Lee1, Shaoling Shang2, and Xiaolong Yu1

1School for the Environment, University of Massachusetts Boston, Boston, MA, USA, 2State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, Fujian, China

Abstract

The Operational Land Imager (OLI) onboard Landsat 8 has potential for mapping the water bio-optical properties with high spatial resolution. Landsat 8/OLI generates the remote sensing reflectance ($R_\text{rs}$) at four visible bands ($\lambda_{1-4}$ = 443, 482, 561, and 655 nm) and is lack of a 412-nm band commonly included for ocean color sensors. This spectral configuration has limited the use of Landsat 8/OLI reflectance product for analytical derivation of light absorption coefficients of phytoplankton, colored dissolved organic matter (CDOM), and detritus. In this study, we proposed a hybrid approach to fill this gap. First, we developed an algorithm to estimate the reflectance in a virtual band centered at $\lambda_0 = 412$ nm from the OLI reflectance spectra $R_\text{rs}(\lambda_{1-4})$. Both the estimated $R_\text{rs}(\lambda_0)$ and measured $R_\text{rs}(\lambda_{1-4})$ were then used together to retrieve the water component absorption coefficients with existing algorithms including the quasi-analytical algorithm. We assessed the model performance using in situ measurements from the global waters. It was found that the proposed approach could estimate $R_\text{rs}(412)$ with a median absolute percentage difference of ~9%. The subsequent retrievals of the component absorption coefficients were satisfactorily accurate, with median absolute percentage difference roughly equal to 35%, 40%, and 60% for phytoplankton, CDOM, and detritus, respectively. The results suggest the feasibility to generate analytically the component absorption coefficients from the Landsat 8/OLI reflectance.

1. Introduction

Phytoplankton, colored dissolved organic matter (CDOM), and detritus are important components of coastal aquatic ecosystems. Phytoplankton harvest spectral radiation to convert inorganic carbon to organic carbon through photosynthesis (Morel, 1991), which sustains the aquatic food web. CDOM and detritus absorb light strongly in the ultraviolet and blue bands, protecting living organisms from damaging radiation (Bricaud et al., 1981). Monitoring the variability of phytoplankton, CDOM, and detritus light absorption coefficients (denoted as $K_{\text{phy}}, K_{\text{cdom}}$, and $K_{\text{det}}$, respectively) is thus crucial for understanding the status of ecological systems, carbon cycles, and water quality related problems (Hansell & Carlson, 2001; Kim et al., 2016; Yu et al., 2016).

Satellite remote sensing represents a unique avenue for synoptic and frequent observation of the global waters. The water color instruments such as the Moderate Resolution Imaging Spectroradiometer and the Visible and Infrared Spectroradiometer can provide the remote sensing reflectance ($R_\text{rs}$, sr$^{-1}$) at multiple bands centered around 412, 443, 488, 550, and 670 nm. These multiband $R_\text{rs}$ data have been extensively used to estimate the phytoplankton, CDOM, and detritus absorption coefficients in various types of waters (Binding et al., 2008; Cao & Miller, 2015; Ciotti & Bricaud, 2006; Mannino et al., 2014; Siegel et al., 2005). However, these satellite data are typical of coarse spatial resolutions (~1 km), insufficient for observation of the bio-optical variability in the dynamic nearshore environments. Landsat 8 satellite is a new member of the land remote sensor family. To the interests of the aquatic science community, the Operational Land Imager (OLI) onboard measures $R_\text{rs}$ with a spatial resolution of ~30 m and with much enhanced radiometric performance at four visible bands ($\lambda_{1-4}$, centered at 443, 482, 561, and 655 nm; Markham et al., 2014). Recent studies show that the OLI reflectance product over waters can be acceptable after atmospheric correction (Pahlevan et al., 2017; Wei et al., 2018). Over the last few years, the OLI observation has quickly found its applications in various aquatic studies. In particular, many empirical methods were used for the estimation of the CDOM absorption coefficient (Alcântara et al., 2016; Olmanson et al., 2016;
Snyder et al., 2017) and phytoplankton chlorophyll a concentration (CHL, milligram per cubic meter; Kim et al., 2016; Lee et al., 2019; Snyder et al., 2017). Still, there exists no operational semianalytical (SA) procedures specific to the OLI data for the derivation of the absorption coefficients from active components, such as phytoplankton. Considering the potential and readiness of the high-resolution satellite observation, it is highly necessary to explore analytical approaches to facilitate the processing and application of the OLI data.

Many SA algorithms exist and can be used to retrieve the absorption coefficient of phytoplankton and the absorption coefficient due to the colored detrital matters (CDM; including CDOM and detritus; denoted as $a_{dg}$) from multiband $R_{rs}$ spectra (Brando et al., 2012; Carder et al., 1999; Hoge & Lyon, 1996; Lee et al., 1998; Lee et al., 2002; Maritorena et al., 2002; Werdell et al., 2013; Werdell et al., 2018). And almost all the SA algorithms require the $R_{rs}(412)$ data in their procedures, because the reflectance at this blue band provides great constraints for robust partition of phytoplankton and CDM (Carder et al., 1991; Wei et al., 2016). For instance, the quasi-analytical algorithm (QAA; Lee et al., 2002) is a stepwise procedure to determine the particle backscattering coefficient ($b_p$) and bulk or total absorption coefficient ($a$), which explicitly uses $R_{rs}$ measurements at 412 nm to partition $a_{ph}$ and $a_{dg}$. The spectral optimization algorithms also specifically use $R_{rs}(412)$ in their procedures (Maritorena et al., 2002; Werdell et al., 2013). The Landsat 8/OLI is lack of a 412-nm band, thus has no $R_{rs}(412)$ measurement. It is challenging to directly apply the existing SA algorithms to the OLI reflectance data for the retrieval of component absorption coefficients.

In this study, we propose a hybrid approach for systematic derivation of the component absorption coefficients from Landsat 8 reflectance data. This approach uses an existing Semi-Analytical algorithm with a Virtual-band Estimator (SAVE). The virtual-band estimator estimates the remote sensing reflectance in a virtual band centered at 412 nm (designated as $\lambda_0$) for an OLI reflectance spectrum $R_{rs}(\lambda_0,4)$. Then $R_{rs}(\lambda_0)$ and $R_{rs}(\lambda_{0,4})$ are combined into a new spectrum, $R_{rs}(\lambda_{0,4})$, which is further used to derive $a$, $a_{ph}$, and $a_{dg}$ products from the QAA algorithm. Last, the $a_{dg}$ product is partitioned into $a_d$ and $a_g$ with an empirical algorithm. The retrievals are validated with in situ measurements from the global waters, which shows promising performance (section 4). As a convenient approach, we recommend the use of SAVE to generate the absorption coefficients of water components from the Landsat 8/OLI reflectance data.

### 2. Algorithm Development and Configuration

#### 2.1. Lookup Table for $R_{rs}$ Spectral Shapes

To estimate $R_{rs}(412)$ from the OLI $R_{rs}(\lambda_{0,4})$ data, a lookup table (LUT) for the $R_{rs}$ spectral shapes centered at $\lambda_{0,4}$ was created from two hyperspectral $R_{rs}$ data sets. The first set of data were in situ measurements (400-800 nm with 3-nm increment) collected from the global waters; a detailed description can be found in Wei et al. (2016). The second data set was simulated with the Hydrolight radiative transfer simulation software (version 5.1; Mobley & Sundman, 2008). For the simulation, we adopted the inherent optical property (IOP) data including the absorption spectra of phytoplankton, detritus, and CDOM and the backscattering spectra of phytoplankton and detritus ($b_{bph}$ and $b_{bd}$) from International Ocean Color Coordinating Group (IOCCG, 2006), with the following modifications. First, we assumed the Fournier-Forand phase function for particle scattering for every model runs, with a constant particle backscattering ratio $b_{bps} = 0.013$ (Whitmire et al., 2007). The pure water absorption coefficient and scattering coefficient were adopted from Lee et al. (2015) and Zhang et al. (2009), respectively. The inelastic scattering was also included in the simulation with default configurations (Mobley & Sundman, 2008), where the chlorophyll fluorescent quantum efficiency was set to 0.02, the CDOM fluorescence was modeled using the spectral fluorescence quantum efficiency function of Mobley (1994), and the Raman scattering cross-section was set to $2.6 \times 10^{-4}$ m$^2$ at the reference wavelength of 488 nm. Two solar zenith angles (30° and 60°) under clear sky were considered for mild sea states (wind speed $w_s = 5$ m/s). The simulation resulted in 1,000 hyperspectral $R_{rs}$ spectra (400-800 nm with 5-nm increment). Collectively, the field data and the synthetic $R_{rs}$ data represent a wide range of waters with CHL varying from ~0.03 to >50 mg/m$^3$. 

---

WEI ET AL. 2
In the second step, the in situ and simulated hyperspectral \( R_{rs} \) spectra were convolved to OLI’s bands \( \lambda_{1-4} \) through

\[
R_{rs}(\lambda_i) = \frac{\int_{400}^{800} R_{rs}(\lambda) RSR_i d\lambda}{\int_{400}^{800} RSR_i d\lambda}, \quad i = 1, 2, 3, \text{ and } 4,
\]

where \( RSR_i \) is the acronym for OLI’s relative spectral responsivity for band \( \lambda_i \). This step is necessary since OLI has wide bandwidths with full width at half maximum of 15, 60, 57, and 37 nm at 443, 482, 561, and 655 nm, respectively. In addition, the hyperspectral \( R_{rs} \) data were interpolated at 412 nm to obtain \( R_{rs}(\lambda_0) \). Here we assumed that the virtual band \( \lambda_0 \) has a narrow bandwidth of 5 nm and hence no spectral weighting was applied to it. The derived five-band \( R_{rs} \) spectra are presented in Figure 1a.

Each individual five-band \( R_{rs} \) spectrum was then normalized by the root of the sum of squares (RSS) of \( R_{rs} \) values from \( \lambda_0 \) to \( \lambda_4 \) as in Wei, Lee, and Shang (2016),

\[
nR_{rs}^*(\lambda_i) = \frac{R_{rs}(\lambda_i)}{\left[ \sum_{i=0}^{4} R_{rs}(\lambda_i) \right]^{1/2}}, \quad i = 0, 1, 2, 3 \text{ and } 4,
\]

where \( nR_{rs}^* \) is the normalized \( R_{rs} \) spectrum. All such normalized spectra are further illustrated in Figure 1b.

From the definition of normalization in equation (2), the \( nR_{rs}^*(\lambda_{0-4}) \) spectra are characteristic of unique features. First, the band ratios of \( nR_{rs}^*(\lambda_{0-4}) \) remain the same with corresponding \( R_{rs}(\lambda_{0-4}) \). Second, the spectral curvature remains unchanged. Third, the magnitudes of \( nR_{rs}^* \) vary from zero to one, inclusively. Last, the RSS of \( nR_{rs}^*(\lambda_{0-4}) \) is always equal to one. These \( nR_{rs}^*(\lambda_{0-4}) \) spectra will be used to represent the \( R_{rs} \) spectral shapes occurring in natural waters perceived by OLI.

### 2.2. Virtual-Band Estimator

With a given Landsat 8/OLI \( R_{rs}(\lambda_{1-4}) \) spectrum, the virtual-band estimator first seeks to identify an \( nR_{rs}^*(\lambda_{0-4}) \) spectrum from the LUT, which has the closest spectral shape to \( R_{rs}(\lambda_{1-4}) \). To do so, we calculate the cosine distance between \( R_{rs}(\lambda_{1-4}) \) and every \( nR_{rs}^* \) spectra in the LUT as

\[
d = 1 - \frac{\sum_{i=1}^{4} [nR_{rs}^*(\lambda_i) \cdot R_{rs}(\lambda_i)]}{\sqrt{\sum_{i=1}^{4} [nR_{rs}^*(\lambda_i)]^2 \sum_{i=1}^{4} [R_{rs}(\lambda_i)]^2}},
\]

where \( d \) is the distance formed by the two vectors of \( nR_{rs}^*(\lambda_{1-4}) \) and \( R_{rs}(\lambda_{1-4}) \). The \( nR_{rs}^* \) spectrum with the minimum cosine distance to \( R_{rs}(\lambda_{1-4}) \) will be determined and chosen for subsequent application.

Figure 1. (a) Remote sensing reflectance spectra at five bands \( \lambda_{0-4} = 412, 443, 482, 561, \) and 655 nm. (b) The lookup table of normalized remote sensing reflectance. As described in the text, \( R_{rs}(\lambda_{1-4}) \) spectra are convolved to Landsat 8/OLI bandwidths.
It is assumed that the “new” spectrum $R_n(\lambda_{0.4})$, composed of known $R_n(\lambda_{1.4})$ and an estimated $R_n(\lambda_0)$, has the spectral shape that best represents the reflectance spectrum of the targeted water. Following equation (2), we normalized the $R_n$ spectrum at 412 nm and further let it be equal to $nR_n'(412)$,

$$\frac{R_n(\lambda_0)}{\left[R_n(\lambda_0)^2 + \sum_{i=1}^{4} R_n(\lambda_i)^2\right]^{1/2}} = nR_n'(412),$$

where the denominator in the left-hand side of equation (4) refers to the RSS of $R_n(\lambda_{0.4})$, while the right-hand side is the value extracted from the above-selected $nR_n'$ spectrum. With only one unknown, equation (4) can be solved for $R_n(\lambda_0)$ as

$$R_n(\lambda_0) = A \cdot nR_n'(412),$$

where $A$ is a scaling factor with

$$A = \left[\sum_{i=1}^{4} (R_n(\lambda_i))^2\right]^{1/2} / \left[\sum_{i=1}^{4} (nR_n'(\lambda_i))^2\right]^{1/2}.$$

With the above estimated $R_n(\lambda_0)$ and known $R_n(\lambda_{1.4})$, the new $R_n(\lambda_{0.4})$ spectrum will be used for the derivation of light absorption coefficients of water components in a procedure described below.

### 2.3. Inversion of Inherent Optical Properties

We adopted the QAA algorithm of Lee et al. (2002) for the IOP inversion from $R_n(\lambda_{0.4})$. The original QAA algorithm has input wavelengths different from OLI’s center wavelengths. In particular, there is a relatively large offset between OLI’s $\lambda_4$ band (655 nm) and QAA’s red band (670 nm). To reduce the uncertainty due to the band mismatch, we converted the input $R_n(\lambda_4)$ to $R_n(670)$ by a polynomial equation ($R^2 = 0.99$),

$$R_n(670) = 10^{P_0 X^3 + P_2 X^2 + P_4 X + P_4},$$

with $X$ the log10-transformed original input $R_n(\lambda_4)$ and the model coefficients $P_0 = 0.0775$, $P_2 = 0.6585$, $P_2 = 2.7692$, and $P_4 = 1.433$. This step is essential for the accurate determination of the total absorption coefficient from QAA when the reference wavelength is set to 670 nm.

QAA first converts the $R_n(\lambda_{0.4})$ spectra to the subsurface remote sensing reflectance ($r_n$) following Lee et al. (2002),

$$r_n(\lambda) = \frac{R_n(\lambda)}{0.52 + 1.7R_n(\lambda)}.$$

Based on numerical simulations of the radiative transfer equations, Gordon et al. (1988) indicated that $r_n$ is a function of the bulk absorption coefficient ($a$) and the total backscattering coefficient ($b_b$),

$$r_n(\lambda) = g_0 b_b(\lambda) + g_1 \left(\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}\right)^2,$$

where $g_0$ and $g_1$ are determined to be 0.089 and 0.125 sr$^{-1}$, respectively (Lee et al., 2002). From the above quadratic equation, $a = b_b/(a + b_b)$ can be solved as a function of $r_n$, $g_0$, and $g_1$.

QAA then proceeds with the estimation of the bulk absorption coefficient at a reference wavelength $a(\lambda_{ref})$ with $\lambda_{ref} = 561$ nm if $R_n(670) < 0.0015$ sr$^{-1}$. Otherwise, $a(\lambda_{ref})$ will be derived at $\lambda_{ref} = 670$ nm. With estimated $a(\lambda_{ref})$, the backscattering coefficient at $\lambda_{ref}$ can be readily derived as below:

$$b_{bp}(\lambda_{ref}) = \frac{u(\lambda_{ref}) x a(\lambda_{ref})}{1 - u(\lambda_{ref})} - b_{bw}(\lambda_{ref}),$$

where $b_{bp}$ is the particle backscattering coefficient and $b_{bw}$ is the backscattering coefficient of pure seawater spectrally weighted by OLI’s RSR function (Lee et al., 2016). From $b_{bp}(\lambda_{ref})$, the spectral backscattering coefficient at bands $\lambda_{0.4}$ can be derived from the power law model (Gordon & Morel, 1983),
\[ b_{lp}(\lambda_i) = b_{lp}(\lambda_{ref}) \left( \frac{\lambda_{ref}^n}{\lambda_i} \right) \eta, \quad i = 0, 1, 2, 3 \text{ and } 4, \]  
\[ a(\lambda_i) = (1-u(\lambda_i))(b_{low}(\lambda_i) + b_{lp}(\lambda_i))/u(\lambda_i). \]  

The QAA algorithm treats the absorption coefficients of CDOM and detritus together, which is analytically solved as below (Lee et al., 2002):

\[ a_{dp}(443) = \frac{|a(412)-\zeta a(443)|-|a_w(412)-\zeta a_w(443)|}{\xi-\zeta}. \]

The parameter \( \zeta \) is estimated as an inverse function of \( r_{rs}(443)/r_{rs}(561) \), while \( \xi \) is equal to \( \exp[S_d(443-412)] \), which is further modeled as an inverse function of \( r_{rs}(443)/r_{rs}(561) \). With determined \( a_{dp}(443) \), the phytoplankton absorption coefficient can be readily derived as the difference between \( a(443) \) and \( a_{dp}(443) \), after accounting for pure seawater’s contribution,

\[ a_{ph}(443) = a(443) - a_{dp}(443) - a_w(443). \]

To further partition \( a_{dp}(443) \) into dissolved and detrital components, we estimated the absorption coefficient of detritus with the algorithm developed by Dong et al. (2013),

\[ a_d(443) = 0.6x^{0.9}, \]

where \( \sigma \) is parameterized as

\[ \sigma = 0.05x a_{pg}(443) + b_{lg}(561) \times 1.4^{R_{rs}(561)-R_{rs}(670)}/R_{rs}(670). \]

This model takes into account the nonwater absorption \( (a_{pg}) \) as well as the backscattering of the particles. The spectral absorption coefficient of \( a_d \) is then quantified as

\[ a_d(\lambda) = a_d(443) \exp[-S_d(\lambda-443)], \]

where the spectral slope \( S_d \) was assumed to be a constant, \( S_d \approx 0.012 \text{ nm}^{-1} \) (Babin et al., 2003). Finally, the absorption coefficient of CDOM was derived as the difference between \( a_{dp}(\lambda) \) and \( a_d(\lambda) \).

The procedures and components of SAVE are schematically shown in a flowchart in Figure 2. The configuration allows for further tuning of the LUT, the virtual-band estimator, and the IOP inversion algorithms whenever necessary in the future.

3. Data and Analysis

3.1. Data Acquisition

We used two independent data sets to assess the accuracy of the SAVE algorithm: the NASA bio-Optical Marine Algorithm Data (NOMAD) (Werdell & Bailey, 2005) and the in situ hyperspectral data retrieved from the SeaWiFS Bio-optical Archive and Storage System (SeaBASS; Hooker et al., 1994). With respect to each data set, we performed the following data reduction.

The NOMAD database consists of multiband spectra for \( R_{rs}, a_{ph}, a_{dp}, a_{pg}, a_d, \) and \( b_{lg} \) centered at 405, 411, 443, 455, 465, 489, 510, 520, 530, 550, 555, 560, 565, 570, 590, 619, 625, 665, 670, and 683 nm. Yet it is noteworthy that many \( R_{rs} \) and IOP values are missing, probably as a result of the multiband instruments used. To create sufficient and usable data for the evaluation in this context, we extended the multiband \( a_d \) and \( a_d \) measurements over a total of 61 spectral bands evenly distributed between 400 and 700 nm following an exponential-decay model as equation (17), with measured \( a_d(443) \) and \( a_d(443) \) and corresponding spectral slope data \( (S_d \text{ and } S_{pg}) \). Similarly, we interpolated linearly the phytoplankton absorption spectra to the
same 61 spectral bands. The hyperspectral \( b_{hp} \) data were obtained by the power law model of equation (11), based on the measured \( b_{hp}(555) \), with the parameter \( \eta \) estimated as \( \eta = \log_{10}(b_{hp}(443)/b_{hp}(510))/\log_{10}(510/443) \). Quality control was further performed to make sure that only the data with 0 < \( \eta \) < 2.5 were kept for subsequent application. With these hyperspectral IOP spectra, we employed Hydrolight to simulate the \( R_{rs} \) spectra. We assumed the same air-water boundary condition, scattering phase function for particles, inelastic scattering, and water depths with the simulations in section 2.1. We only considered a clear sky with the solar zenith angle at 30°. As a result, we obtained 454 hyperspectral \( R_{rs} \) spectra for the given IOP data.

The SeaBASS database archives a large amount of \( R_{rs} \) measurements with IOP data. For the present analyses, we extracted the hyperspectral data from selected nearshore waters including Boston Harbor and Massachusetts Bay (2017-2018), the Great Lakes (2013-2014), Chesapeake Bay (2011), and the northern Gulf of Mexico (2013). The majority of the \( R_{rs} \) spectra were measured by the HyperPro radiometers floating at the water surface with a skylight-blocking approach (137 bands, 350-800; Lee et al., 2013) or profiling of the upper water columns (Mueller et al., 2003). In Chesapeake Bay, the radiometric data were measured with a hand-held radiometer and postprocessed using the method of Lee et al. (2010). The CDOM

Figure 2. Flowchart for the Semi-Analytical algorithm with a Virtual-band Estimator algorithm to estimate phytoplankton, colored dissolved organic matter, and detritus absorption coefficients from Landsat 8/OLI \( R_{rs}(\lambda_{1-4}) \) spectra.
samples were collected through 0.2-μm filters, and the CDOM absorption coefficient was measured by the scanning spectrophotometers (250-800 nm). Samples for particulate absorption were captured using 25 mm precombusted GF/F filters (pore size 0.7 μm), and the absorption coefficient of particles \( (a_p) \) was determined using the quantitative filter pad technique (QFT; Mitchell et al., 2002). The particles-loaded filter was later bleached using methanol to remove phytoplankton pigments to determine the absorption coefficient by nonalgae particles. The phytoplankton absorption coefficient was then derived as the difference between \( a_p \) and \( a_d \) (\( a_{ph} = a_p - a_d \)).

The hyperspectral \( R_{rs} \) and backscattering data were convolved to OLI’s bands at \( \lambda_{1-4} \) according to equation (1). The absorption coefficient is inversely proportional to reflectance and was derived following Lee et al. (2016),

\[
a_x(\lambda_i) = \frac{\int_{400}^{800} RSR_i d\lambda}{\int_{400}^{800} 1/[a_x(\lambda)] RSR_i d\lambda} \quad i = 1, 2, 3, \text{ and } 4,
\]

where \( a_x \) refers to \( a_{ph}, a_{pg}, a_g \), or \( a_d \). At \( \lambda_0 = 412 \) nm, the corresponding \( R_{rs} \) values were interpolated from the hyperspectral data, without spectral convolution.

3.2. Validation Analyses

To quantify the accuracy of model-estimated quantities, the median absolute percentage difference (MAPD) was derived as

\[
MAPD = \text{median}\{ \frac{|(M_j - T_j)|}{T_j} \times 100\% \},
\]

where \( M_j \) and \( T_j \) refer to the estimated and known values under investigation, respectively, and \( j \) varies from 1 to \( n \), the total number of valid observation involved in the evaluation. The signed relative difference or error (\( \delta \)) was calculated for the model-estimated \( R_{rs}(412) \) values, as the following:

\[
\delta = \frac{(M_j - T_j)}{T_j} \times 100\%.
\]

Besides, the root mean square difference (RMSD) was analyzed for each estimated and known quantities, defined as

\[
RMSD = \left[ \frac{1}{n-1} \sum_{j=1}^{n} \left( \frac{M_j - T_j}{T_j} \right)^2 \right]^{1/2}.
\]

We also performed the model II regression analyses of the log10-transformed quantities and computed the slope and the coefficient of determination (\( R^2 \)). Nonrealistic retrievals such as the negative values due to imperfect model architectures were excluded from the statistical analysis. As a result, the number of valid retrievals involved in the validation analyses, \( n \), can sometimes be less than the total number of input \( R_{rs} \) spectra, \( N \).

3.3. Sensitivity Analyses

To understand how the uncertainty of the estimated \( R_{rs}(412) \) impacts the retrieved absorption coefficients, we carried out the following sensitivity analyses. The IOCCG synthetic data (IOCCG, 2006), considered free of measurement uncertainty, were used. The data (\( N = 500 \)) were divided into three subgroups with respect to their chlorophyll concentrations: the eutrophic \( (\text{CHL} \geq 1 \text{ mg/m}^3) \), mesotrophic \( (0.1 \leq \text{CHL} < 1 \text{ mg/m}^3) \), and oligotrophic waters \( (\text{CHL} < 0.1 \text{ mg/m}^3) \). We convolved the \( R_{rs} \) and IOP spectra of each subgroup to Landsat 8/OLI bands and interpolated them to \( \lambda_0 \). Assuming the “error-free” \( R_{rs}(\lambda_{1-4}) \) spectra, we disturbed the \( R_{rs}(412) \) values with random errors,

\[
R_{rs}^{err}(412) = R_{rs}(412) + \Re \cdot \delta,
\]

where \( R_{rs}^{err}(412) \) represents the error-disturbed values, \( \Re \) is a random number between 0 and 1 of standard normal distribution, and \( \delta \) is the relative difference of \( R_{rs}(412) \) to its true value. Nine instances of \( \delta \) were
considered, which vary from \(-20\%\), \(-15\%\), \(-10\%\), \(-5\%\), 0\%, +5\%, +10\%, +15\%, and finally to +20\%. The SAVE algorithm was implemented directly with these erroneous Rrs spectra following the work flow starting with equation (7) (Figure 2), that is, revoking the virtual-band estimator. The median of the unbiased percentage difference (denoted as \(\Delta\)) was calculated between the model-estimated absorption coefficients from the erroneous Rrs spectra and those from error-free Rrs spectra, that is, \(\delta R_{\text{rs}}(412) = 0\), as

\[
\Delta = \text{median} \left\{ 2 \times \left[ \frac{E_j - F_j}{E_j + F_j} \right] \times 100\% \right\}, \quad j = 1, 2, ..., n,
\]

where \(E_j\) refers to the model retrievals with the error-disturbed Rrs spectra, while \(F_j\) is the corresponding optical properties inverted from the error-free Rrs spectra. In the following, we treat \(\Delta\) as the uncertainty of the model-estimated absorption coefficients originated from \(\delta R_{\text{rs}}(412)\).

4. Results

4.1. Evaluation of \(R_{\text{rs}}(412)\) Retrievals

As \(R_{\text{rs}}(412)\) is critical for the proposed algorithm, we first evaluate the accuracy for the estimated \(R_{\text{rs}}(412)\) values (see Figure 3a and Table 1). For the NOMAD data, the estimated \(R_{\text{rs}}(412)\) are reasonably consistent with known values with small differences (MAPD = 7\% and RMSD = 0.00044 sr\(^{-1}\)). The SeaBASS data span a relatively narrower dynamic range, with \(R_{\text{rs}}(412)\) varying between 0.001 and 0.009 sr\(^{-1}\); the estimated \(R_{\text{rs}}(412)\) data are overall consistent with the measured values of SeaBASS data (MAPD = 11\% and RMSD = 0.00045 sr\(^{-1}\)). In addition, regression analyses between the estimated and known \(R_{\text{rs}}(412)\) values obtained linear slopes only slightly deviated from unity (1.05 and 1.04 for NOMAD and SeaBASS data, respectively). Furthermore, we compared the frequency distribution of the relative errors of the estimated \(R_{\text{rs}}(412)\) (Figure 3b). It is found that the relative errors (\(\delta R_{\text{rs}}(412)\)) are positively skewed. The NOMAD estimation has the first mode at 2\%, while the SeaBASS estimation shows a larger first mode at around 12\%. This discrepancy can be partly attributable to the uncertainty or errors of the in situ \(R_{\text{rs}}\) measurements and partly to the uncertainty of the algorithm itself. In spite of the biases, more than 90\% of the estimated \(R_{\text{rs}}(412)\) are found subjected to a relatively small error with \(\delta R_{\text{rs}}(412)\) varying between \(-20\%\) and 20\% for the SeaBASS data. A few outliers are present with the estimated SeaBASS \(R_{\text{rs}}(412)\) (Figure 3a), which contribute to the positive tail of the frequency distribution (Figure 3b). These outliers represent highly absorptive waters, with \(R_{\text{rs}}(412) < 0.001\) sr\(^{-1}\). The measurement uncertainties of these \(R_{\text{rs}}\) spectra are partly responsible for the discrepancy between estimated and known \(R_{\text{rs}}(412)\) values.
The variation of $\delta R_{\text{rs}}(412)$ was further assessed with respect to the chlorophyll concentrations (Figure 4a). Interestingly, there is a slightly negative trend between $\delta R_{\text{rs}}(412)$ and CHL. The $\delta R_{\text{rs}}(412)$ slowly vary from positive to negative when CHL increases from <0.1 to >10 mg/m$^3$. On the other hand, a weak but positive trend exists between $\delta R_{\text{rs}}(412)$ and $R_{\text{rs}}(412)$ (Figure 4b). The dependencies of $\delta R_{\text{rs}}(412)$ on CHL and $R_{\text{rs}}(412)$ are not contradictory, as smaller $R_{\text{rs}}(412)$ values are often suggestive of stronger phytoplankton and CDOM absorption, and vice versa. According to the illustrations in Figure 4, the $\delta R_{\text{rs}}(412)$ can be relatively larger in very low-CHL waters or very high-CHL waters. For the former case, the $R_{\text{rs}}(412)$ values are high, but the $R_{\text{rs}}(\lambda_{1-4})$ spectra usually vary over a narrow range, which may impact the spectral-shape-based classification as equation (3). For the latter case, the $R_{\text{rs}}(412)$ values can be very small, leading to larger percentage differences.

Overall, the differences between the estimated and known $R_{\text{rs}}(412)$ values are small. According to the validation analyses, the satellite-derived $R_{\text{rs}}(412)$ are often subject to an MAPD of greater than 20% from their in situ matchups in nearshore environments (Qin et al., 2017; Zibordi et al., 2009). In addition, the in situ measurements of $R_{\text{rs}}(412)$ from different measurement platforms can differ from each other with MAPD > ~10% (Hooker et al., 2002). So we can conclude that the above-discussed small differences for the estimated $R_{\text{rs}}(412)$ provide a confident measure of the virtual-band estimator in these waters.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>NOMAD data (N = 454)</th>
<th>SeaBASS data (N = 266)</th>
<th>Rrs(412)</th>
<th>Slope</th>
<th>R²</th>
<th>MAPD</th>
<th>RMSDa</th>
<th>n</th>
<th>Slope</th>
<th>R²</th>
<th>MAPD</th>
<th>RMSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta R_{\text{rs}}$</td>
<td>412</td>
<td>454</td>
<td>1.05</td>
<td>0.99</td>
<td>7%</td>
<td>0.00044</td>
<td>266</td>
<td>1.04</td>
<td>0.94</td>
<td>11%</td>
<td>0.00045</td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>412</td>
<td>454</td>
<td>0.92</td>
<td>0.97</td>
<td>19%</td>
<td>0.27</td>
<td>266</td>
<td>1.01</td>
<td>0.90</td>
<td>17%</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>$a_{\text{ph}}$</td>
<td>412</td>
<td>442</td>
<td>0.90</td>
<td>0.76</td>
<td>33%</td>
<td>0.13</td>
<td>255</td>
<td>0.97</td>
<td>0.68</td>
<td>38%</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>$a_{\text{dg}}$</td>
<td>412</td>
<td>454</td>
<td>0.97</td>
<td>0.93</td>
<td>28%</td>
<td>0.28</td>
<td>266</td>
<td>0.97</td>
<td>0.86</td>
<td>25%</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>$a_{g}$</td>
<td>412</td>
<td>454</td>
<td>0.97</td>
<td>0.92</td>
<td>30%</td>
<td>0.18</td>
<td>266</td>
<td>0.97</td>
<td>0.78</td>
<td>32%</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>$a_{d}$</td>
<td>412</td>
<td>454</td>
<td>0.97</td>
<td>0.81</td>
<td>67%</td>
<td>0.21</td>
<td>266</td>
<td>0.97</td>
<td>0.86</td>
<td>25%</td>
<td>0.44</td>
<td></td>
</tr>
</tbody>
</table>

Note. SAVE = SemiAnalytical algorithm with a Virtual-band Estimator; NOMAD = NASA bio-Optical Marine Algorithm Data; SeaBASS = SeaWiFS Bio-optical Archive and Storage System; MAPD = median absolute percentage difference; RMSD = root mean square difference. Note that the regression statistics was performed with log-transformed data. N is the number of $R_{\text{rs}}$ data tested, and n is the number of valid model retrievals.

*The unit is sr$^{-1}$ for $R_{\text{rs}}$ data and m$^{-1}$ for absorption coefficient.*

The variation of $\delta R_{\text{rs}}(412)$ was further assessed with respect to the chlorophyll concentrations (Figure 4a). Interestingly, there is a slightly negative trend between $\delta R_{\text{rs}}(412)$ and CHL. The $\delta R_{\text{rs}}(412)$ slowly vary from positive to negative when CHL increases from <0.1 to >10 mg/m$^3$. On the other hand, a weak but positive trend exists between $\delta R_{\text{rs}}(412)$ and $R_{\text{rs}}(412)$ (Figure 4b). The dependencies of $\delta R_{\text{rs}}(412)$ on CHL and $R_{\text{rs}}(412)$ are not contradictory, as smaller $R_{\text{rs}}(412)$ values are often suggestive of stronger phytoplankton and CDOM absorption, and vice versa. According to the illustrations in Figure 4, the $\delta R_{\text{rs}}(412)$ can be relatively larger in very low-CHL waters or very high-CHL waters. For the former case, the $R_{\text{rs}}(412)$ values are high, but the $R_{\text{rs}}(\lambda_{1-4})$ spectra usually vary over a narrow range, which may impact the spectral-shape-based classification as equation (3). For the latter case, the $R_{\text{rs}}(412)$ values can be very small, leading to larger percentage differences.

Overall, the differences between the estimated and known $R_{\text{rs}}(412)$ values are small. According to the validation analyses, the satellite-derived $R_{\text{rs}}(412)$ are often subject to an MAPD of greater than 20% from their in situ matchups in nearshore environments (Qin et al., 2017; Zibordi et al., 2009). In addition, the in situ measurements of $R_{\text{rs}}(412)$ from different measurement platforms can differ from each other with MAPD > ~10% (Hooker et al., 2002). So we can conclude that the above-discussed small differences for the estimated $R_{\text{rs}}(412)$ provide a confident measure of the virtual-band estimator in these waters.

![Figure 4](image-url). Variation of $\delta R_{\text{rs}}(412)$ with water trophic status (panel a) and the magnitudes of $R_{\text{rs}}(412)$ (panel b) for NASA bio-Optical Marine Algorithm Data (NOMAD; denoted by open circles) and SeaWiFS Bio-optical Archive and Storage System (SeaBASS) data (denoted by crosses). The legends are the same as Figure 3a. Note that the number of in situ data used in (a) and (b) is different due to data availability.
4.2. Evaluation of Bulk, Phytoplankton, and CDM Absorption Coefficient Retrievals

The model-estimated bulk absorption coefficients at 412 and 443 nm are favorably comparable to known values (Figure 5 and Table 1). First, we stress that the retrieval of $a_{(443)}$ is irrelevant to the estimation of $R_{rs}(412)$ under the framework of QAA. Rather, $a_{(443)}$ is solely determined by $R_{rs}(\lambda_{1-4})$. For NOMAD, the $a_{(412)}$ and $a_{(443)}$ retrievals agree well (MAPD = 16-17%, RMSD = 0.22-0.2 m$^{-1}$, and $R^2 = 0.97$) with known values over the full dynamic range of the test data (0.02-3 m$^{-1}$). The retrievals of $a_{(412)}$ and $a_{(443)}$ for SeaBASS are slightly less accurate (MAPD = 17-19%, RMSD = 0.36-0.4 m$^{-1}$, and $R^2 = 0.86-0.9$), a result of the uncertainties from both $R_{rs}$ and IOP measurements. The retrievals for the component absorption coefficients $a_{ph}$ and $a_{dg}$ suffer larger uncertainties than the total absorption coefficients (Table 1 and Figure 6). This is related to the step-by-step nature of ocean color inversion as implemented by QAA where $R_{rs}$ is governed by the bulk optical properties, not individual components. Therefore more uncertainties or errors are expected in the retrieved component absorption coefficients $a_{ph}(443)$ and $a_{dg}(443)$ when they are partitioned from $a_{(412)}$ and $a_{(443)}$ (Figure 2). The uncertainties in $a_{(412)}$ and $a_{(443)}$ propagate to and impact the estimated $a_{ph}(443)$ and $a_{dg}(443)$ (Lee et al., 2010). Considering the error metrics and regression analyses, the agreement between retrieved and known $a_{dg}$ values at two blue bands is better than $a_{ph}$. The differential performance for $a_{dg}$ and $a_{ph}$ estimation is common with the SA algorithms (IOCCG, 2006; Werdell et al., 2013) and can be partly explained by the fact that $a_{dg}$ dominates the total absorption coefficients at two blue bands for the majority (>80%) of the NOMAD and SeaBASS data used herein. In addition, it is also likely related to the difficulty in modeling the spectral shapes for phytoplankton absorption, which vary significantly in natural waters (Bricaud et al., 2004; Roesler et al., 1989). Some outliers are present with the SeaBASS $a_{dg}$ retrievals, which are most probably a result of the problems with sampling of inhomogeneous waters (IOCCG, 2018). The uncertainties induced by these outliers further propagated to the total absorption coefficients and affected their comparison (Figure 5). In comparison, the retrievals for SeaBASS data have shown larger variability (with higher RMSD) than those for NOMAD data. This is partly related to the fact that we simulated the $R_{rs}$ spectra for the NOMAD data, which are free of measurement uncertainty.

The performance of our approach in estimating $a_{(412)}$ is comparable to other SA algorithms, which used the $R_{rs}(412)$ measurements for the inversion (e.g., Loisel et al., 2018; Werdell et al., 2013). The MAPD values varying between 30% and 39% (see Table 1) are not a small uncertainty for $a_{ph}(443)$ and $a_{dg}(443)$ estimation. But it is not uncommon to observe such evaluation results for ocean color inversion algorithms, since the in situ data and the SA algorithms are always subject to uncertainties. For comparison, Werdell et al. (2013) showed that the estimated NOMAD $a_{ph}(443)$ and $a_{dg}(443)$ from the generalized IOP (GIOP) algorithm have an uncertainty with MAPD equal to 26% and 35%, respectively; their results compare favorably to ours (30% and 30% for $a_{ph}(443)$ and $a_{dg}(443)$, respectively). Note that they used more data points from the NOMAD data set ($n = 682$) than the present study ($n = 454$) but the absorption coefficients span approximately the
same dynamic range as the present analyses. Smyth et al. (2006) analyzed their SA algorithm with NOMAD data \((n = 459)\) and reported for the absorption retrievals less accurate than the results shown here, which is partly related to the narrower dynamic range of the data used therein. From the above analyses, we can reach a conclusion that our approach can generate retrievals for the bulk, phytoplankton, and CDM absorption coefficients from Landsat 8/OLI reflectance, with accuracy comparable to previous analyses.

### 4.3. Evaluation of CDOM and Detritus Absorption Coefficient Retrievals

It is important to reiterate that the SA algorithms generally do not partition the absorption coefficients of CDOM and detritus because of their similar spectral behaviors (Lee et al., 2002; Maritorena et al., 2002; Smyth et al., 2006; Werdell et al., 2013). The procedure adopted herein for the estimation of detritus absorption \(a_d\) is empirical in nature. It suffers the uncertainties from the estimation of \(a(443)\) and \(b_{pp}(561)\) but is irrelevant to \(R_{rs}(412)\). As the intermediate product, the \(a_d\) retrievals at 412 and 443 nm are subject to large MAPD varying between 56% and 67%, in spite of the high \(R^2\) and close-to-one slopes (Figure 7 and Table 1). The \(a_d(412)\) and \(a_d(443)\) retrievals tend to be overestimated for \(a_d < 0.01\) m\(^{-1}\) and underestimated for \(a_d > 0.1\) m\(^{-1}\). The negative bias is particularly significant for the NOMAD data when \(a_d > 0.1\) m\(^{-1}\). This may be caused by the measurement uncertainty of the particle absorption coefficients (IOC, 2018; Neeley et al., 2015). In contrast, the retrieval of the CDOM absorption coefficient \(a_g\) relies on the

![Figure 6. Comparison of Semi-Analytical algorithm with a Virtual-band Estimator (SAVE)-estimated \(a_{ph}\) and \(a_{dg}\) with known values for the NASA bio-Optical Marine Algorithm Data (NOMAD; denote by open circles) and SeaWiFS Bio-optical Archive and Storage System (SeaBASS) data (denoted by crosses). The regression and validation statistics is given in Table 1.](image-url)
knowledge of $R_{rs}(412)$, which allows for the derivation of $a_g$ under the framework of QAA. The estimated $a_g(443)$ are much more accurate than $a_d(443)$. This is mostly because the values of $a_d(443)$ are smaller than $a_g(443)$ for the majority of the NOMAD data (>90%) and SeaBASS data (>70%). Our comparison indicates that the accuracy of $a_g(412)$ retrievals are higher than or at least equivalent to $a_g(443)$. This is valuable as the CDOM absorption is strongly correlated with the dissolved organic carbon and salinity in coastal waters (Del Castillo & Miller, 2008; Pan & Wong, 2015; Vantrepotte et al., 2015). Undoubtedly, the SA derivation of $a_g(412)$ (and likely $a_g(443)$) with the present approach will provide support to the study of the carbon stocks and material transport in nearshore environments.

### 4.4. Sensitivity of Absorption Retrievals to $R_{rs}(412)$ Estimation

The sensitivity of the absorption coefficient retrievals to the relative error of $R_{rs}(412)$ ($\delta R_{rs}(412)$) is summarized in Table 2. Because the estimation of $a(443)$ from QAA does not rely on $R_{rs}(412)$, the uncertainty for $a(443)$ retrieval ($\Delta a(443)$) is independent of $\delta R_{rs}(412)$. As anticipated, however, the uncertainties of other absorption coefficient retrievals do vary with $\delta R_{rs}(412)$. The $a(412)$ values are estimated from $R_{rs}(412)$ and $b_{pp}(412)$; the latter is irrelevant to the estimation of $R_{rs}(412)$ but determined by the $R_{rs}(443)/R_{rs}(561)$ ratios (refer to the QAA algorithm; Lee et al., 2002). So it is straightforward to understand that $\Delta a(412)$ increases...
with increasing $\delta R_{rs}(412)$. Because we have assumed that $R_{rs}(443)$ and $R_{rs}(561)$ are error free in our analysis, the rate of change in $\Delta R_{rs}(412)$ is found weak, with resultant $\Delta a(412)$ values smaller than the absolute amplitudes of $\Delta R_{rs}(412)$. The uncertainties of the component absorption coefficients are much more sensitive to $\delta R_{rs}(412)$, where $\Delta a_{ph}$ and $\Delta a_g$ can be twice larger than the absolute magnitudes of $\Delta R_{rs}(412)$. In the mesotrophic waters, for instance, $\Delta a_{ph}(412)$ and $\Delta a_g(443)$ increase from <\~27% to <\~45% when $\delta R_{rs}(412)$ varies from ±10% to ±20%. This result can be explained by the involvement of $a(412)$ in the separation of $a_{ph}(443)$ and $a_g(443)$, which plays a major role in $\Delta a_{ph}(412)$ and $\Delta a_g(443)$ in QAA (Lee, Arnone, et al., 2010). Among all component absorptions, $\Delta a_{dg}$ remains least sensitive to $\delta R_{rs}(412)$. In all three types of waters, $\Delta a_{dg}$ and $\Delta a_g$ are found roughly comparable to each other. Our analyses only indicate a weak dependence of $\Delta a_{ph}$ and $\Delta a_{dg}$ on the chlorophyll a concentrations. Specifically, $\Delta a_{ph}(412)$ and $\Delta a_g(443)$ remain the smallest in the oligotrophic waters, while $\Delta a_{dg}(412)$ and $\Delta a_g(443)$ are the smallest for the eutrophic and mesotrophic waters. For CDOM, we found the smallest $\Delta a_g$ in the mesotrophic waters. Unlike all above component absorption coefficients, $\Delta a_d$ does not change with $\delta R_{rs}(412)$, because its derivation is not related to $R_{rs}(412)$ in the present configuration (recalling equations (15) and (16)).

### 5. Discussion

The Landsat 8/OLI instrument has the potential for generating high-spatial-resolution bio-optical properties in the dynamic nearshore waters. Compared with the operational satellite sensors such as Moderate Resolution Imaging Spectroradiometer and Visible and Infrared Spectroradiometer, Landsat 8/OLI is still short-handed because of the relatively fewer number of visible bands ($\lambda = 443, 482, 561$, and 655 nm). Particularly, the lack of a blue band at 412 nm has made it difficult to implement the existing SA algorithms (IOCCG, 2006; Lee et al., 2002; Maritorena et al., 2002; Smyth et al., 2006; Werdell et al., 2013) for the retrieval of component absorption coefficients. The SAVE algorithm is designed to fill this gap. It predicts $R_{rs}$ in a virtual band centered at $\lambda_0 = 412$ nm for each individual $R_{rs}(\lambda)$ spectrum. The estimated $R_{rs}(\lambda)$ and existing $R_{rs}(\lambda_{1-4})$ together allow for analytical derivation of absorption coefficients of the major components. This convenient approach makes it practical for generating the absorption coefficients for phytoplankton, CDOM, and detritus from the Landsat 8/OLI data while following the framework of the existing SA algorithms.

The virtual-band estimator is a novel and critical component of the SAVE algorithm. In spite of the encouraging performance, the model still faces challenges. First, the virtual-band estimator evokes spectral matching (or optical classification) based on four-band $R_{rs}$ spectra (recall equation (3)). With too few bands, this matching can introduce uncertainty or error, albeit small. Second, it is impossible for the lookup table to include every $R_{rs}$ spectral shapes existing in natural waters. When the actual $R_{rs}$ shape deviates from the LUT, the matching and subsequent estimation of $R_{rs}(412)$ may suffer sometimes large uncertainty or error. Nevertheless, the analyses suggest that the SAVE algorithm provides a satisfactory estimation for $R_{rs}(412)$ over the high dynamic range of waters with $R_{rs}(412)$ varying between <\~0.001 and 0.025 sr$^{-1}$. The study here used the NOMAD and SeaBASS data, which are not free of measurement uncertainty. The uncertainty for the particulate absorption coefficient measurements can sometimes be very large (Neely et al., 2015). The in situ $R_{rs}$ measurements are also subject to uncertainties originating from the calibration, environmental disturbance, postprocessing, and so forth (Wei et al., 2014; Zibordi et al., 2012).

### Table 2

The Uncertainty Δ ($\times100$) of SAVE-Estimated Absorption Coefficients

<table>
<thead>
<tr>
<th>$\delta R_{rs}$</th>
<th>$\Delta \Delta \Delta \delta \delta \delta \delta \delta$</th>
<th>$\Delta \Delta \Delta \delta \delta \delta \delta \delta$</th>
<th>$\Delta \Delta \Delta \delta \delta \delta \delta \delta$</th>
<th>$\Delta \Delta \Delta \delta \delta \delta \delta \delta$</th>
<th>$\Delta \Delta \Delta \delta \delta \delta \delta \delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta a_{ph}$</td>
<td>412</td>
<td>14</td>
<td>-10%</td>
<td>-10%</td>
<td>-10%</td>
</tr>
<tr>
<td></td>
<td>443</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\Delta a_{dg}$</td>
<td>412</td>
<td>44</td>
<td>34</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>443</td>
<td>44</td>
<td>34</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>$\Delta a_g$</td>
<td>412</td>
<td>42</td>
<td>29</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>443</td>
<td>42</td>
<td>29</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

Note. SAVE = Semi-Analytical algorithm with a Virtual-band Estimator.
difficult to determine, these measurement uncertainties have certainly affected the validation results. According to the sensitivity analyses (Table 2), the influence of the erroneous $R_{ns}(412)$ input data is indeed not negligible under the extreme conditions when $\delta R_{ns}(412)$ approach $\pm 20\%$. The assumption of $\pm 20\%$ errors for $R_{ns}(412)$ is not randomly chosen but inferred from the evaluations of the in situ measurements (Figure 3). We did not further change $\delta R_{ns}(412)$, positively or negatively, because the resultant $R_{ns}$ spectral shapes could be dramatically distorted so that the resultant $R_{ns}$ spectra become unrealistic. A complete understanding of the model uncertainty will demand an investigation of the uncertainty propagation with globally representative data.

As our analyses demonstrated, SAVE provides a feasible avenue for the analytical retrieval of the water component absorption coefficients, particularly for phytoplankton, CDM, and CDOM from the Landsat 8/OLI reflectance data for global applications (Tables 1 and 2). The success of the model benefits from the use of existing SA algorithm, specifically, of QAA (Lee et al., 2002). It is interesting to note the model configuration also allows utilizing the other SA algorithms for optical inversion. For example, one may replace QAA with the spectral-optimization algorithm of GIOP (Werdell et al., 2013) for the purpose of the IOP inversion. Being a temporary solution, the current modeling framework does not rule out the necessity of continuous development of the SA schemes to further improve the model performance for the derivation of light absorption coefficients.

To demonstrate the application of the SAVE algorithm for the monitoring of nearshore waters, we derived the absorption coefficient products from a Landsat 8/OLI image. The Landsat 8/OLI image was captured on 28 August 2015 over Boston Harbor and Massachusetts Bay (Figure 8). The OLI bands 5 and 7 (865 and 2201 nm, respectively) were used for the atmospheric correction and the satellite retrievals flagged as ATMPFAIL, LAND, CLDICE, and HILT were masked out. The validation of Landsat 8/OLI $R_{ns}$ spectra with in situ matchups were provided elsewhere (Wei et al., 2018). As shown in Figure 8a, the transport pattern of surface water including the water plumes, bright/dark waterfronts, and ship wakes are captured in the composited true color image. The spatial gradients and distribution patterns of the absorption coefficients for CDM, phytoplankton, and CDOM from the Harbor toward the Bay are clearly shown up in Figures 8b–8d. Within the spatial domain under investigation, the CDOM and the detritus dominate the light absorption coefficients instead of phytoplankton. Also, $a_{ph}(443)$, $a_{dg}(443)$, and $a_{ag}(443)$ are found extremely high near the shorelines and river mouths with values greater than 10 m$^{-1}$ for Landsat 8/OLI measurements of this time. The absorption products in the strongly absorptive waters with $a > 1$ m$^{-1}$ have not been fully validated though due to lack of concurrent field measurements. As suggested in Figures 8b–8d, there exists significant difference in the distribution patterns of $a_{ph}(443)$ and $a_{dg}(443)$ (and $a_{ag}(443)$) products. Such difference is attributable to the fact that the phytoplankton does not covary with CDM, CDOM, or detritus, a phenomenon commonly occurring in optically complex nearshore waters. The decoupling of $a_{ph}(443)$ and $a_{ag}(443)$ is of significant implication for the estimation of phytoplankton biomass in such aquatic environments. In water color remote sensing, the empirical blue-green spectral band ratios of $R_{ns}$ spectra such as $R_{ns}(482)/R_{ns}(561)$ are often used for CHL retrievals (Franz et al., 2015; Snyder et al., 2017). The band-ratio algorithms are sensitive to the bulk absorption and work relatively well for “case 1” water (Gregg & Casey, 2004; Moore et al., 2009). But the reflectance ratios become less sensitive to the changes of CHL as the CDOM and/or detritus absorption starts to play increasingly important roles in the bulk absorption (Dierssen, 2010). As a result, the comparison of the CHL retrieval from the band-ratio algorithm (Figure 8e) with the $a_{ph}(443)$ retrieval can show strikingly different spatial distribution patterns. For these particular observations, the spatial distribution of the CHL product is much more similar to $a_{dg}(443)$ and $a_{ag}(443)$, indicating erroneous CHL product for such nearshore waters when it is estimated using blue-green band-ratio algorithms. The SA retrieval of $a_{ph}(443)$ from the current approach likely provides more accurate estimation for the autotrophic biomass with fine details.

The framework of SAVE is designed with the objective to facilitate the use of the Landsat 8/OLI reflectance data. It is certainly feasible to implement it with reflectance measurements from other radiometers as long as they share similar band settings with the OLI sensor. For instance, the Sentinel 2 satellite is equipped with the MultiSpectral Instrument (MSI). The MSI has four visible bands centered at 444, 497, 560, and 664 nm, almost identical to Landsat 8/OLI. The future Landsat 9 satellite will have the OLI-2. The algorithm developed in this study should be readily applicable to MSI and OLI-2 imageries.
6. Concluding Remarks

Landsat 8/OLI is an advanced multiband and high-resolution satellite sensor and can generate high-quality reflectance product for the coastal waters. Since its launch in 2013, the Landsat 8 data have been widely used in water-related studies on a variety of topics. Partly due to the lack of a 412-nm band, however, it remains challenging to apply existing SA algorithms to Landsat 8/OLI reflectance product for derivation of the light absorption coefficients for various water components. Consequently, the Landsat 8 data have not received sufficient attention in wider applications to coastal water remote sensing as it should have deserved after...
6-year operation. In this study, we developed SAVE to facilitate the derivation of phytoplankton, CDOM, and detritus absorption coefficients from the Landsat 8/OLI reflectance. SAVE employs a virtual-band estimator to estimate $R_n$ at 412 nm from the existing reflectance spectra measured at other bands. It is found that such estimated $R_n(412)$ is acceptably accurate over the full dynamic range of the test data, with MAPD of 7% and 11% for the NOMAD and SeaBASS data, respectively. Thus, the estimation of $R_n(412)$ allows to use existing SA inversion algorithms for derivation of the component absorption coefficients of waters. In current analysis, we adopted the QAA to derive $a_{ph}(443)$ and $a_{dg}(443)$. The $a_{ph}(443)$ retrievals are found only subject to an uncertainty with MAPD = 30% and MAPD = 39% for NOMAD and SeaBASS data, respectively. The estimated $a_{dg}(443)$ suffer an uncertainty with MAPD = 30% for NOMAD and MAPD = 32% for SeaBASS data. These uncertainty statistics are comparable with those of existing SA algorithms using the $R_n(412)$ measurements. The SAVE algorithm further partitions $a_{dg}$ into $a_{p}$ and $a_{d}$ with an existing algorithm of Dong et al. (2013). The accuracy of the estimated $a_{p}$ (MAPD = 35–47%) is found comparable with $a_{ph}$ retrievals. Our analyses suggest that it is feasible to use the new approach to semianalytically generate the component absorption coefficients for the global waters with $a(443)$ varying between 0.01 and 3 m$^{-1}$, with acceptable accuracy. The semianalytically derived phytoplankton absorption coefficient is arguably a more reliable proxy for phytoplankton biomass in coastal waters where the phytoplankton is often decoupled with CDOM and detritus. On the other hand, the CDOM and detritus absorption coefficients will also provide important information for aquatic biology, carbon cycles, and other climate-related problems.

Acknowledgments
This study was supported by the NASA projects (NNX15AC84G, NNX16AD38G, and 80NSSC18K0509), the NOAA VIIRS Cal/Val project (NA11OAR4320199), and the Natural Science Foundation of China (NSFC) (No.41576169). We thank the NASA OBPG for distributing the NOMAD and SeaBASS data and the principal investigators for contributing data to SeaBASS. Thanks also go to two anonymous reviewers for comments and recommendation. The MATLAB script is developed for the current algorithm and accessible online (http://oceanoptics.umb.edu/resources/). Data are available from authors.

References
JRS.9.096070


