This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0
https://creativecommons.org/publicdomain/mark/1.0/

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Research Article Vol. 62, No. 13 / 1 May 2023 / Applied Optics 3299

applied optics

Optimizing retrieval spaces of bio-optical models for remote sensing of ocean color

Neranga K. Hannadige,[1] Peng-Wang Zhai,[1,*] P. Jeremy Werdell,[2] Meng Gao,[2,3] Bryan A. Franz,[2] Kirk Knobelspiesse,[2] and Amir Ibrahim[2]

[1]Joint Center for Earth Systems Technology, Department of Physics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250, USA
[2]NASA Goddard Space Flight Center, Code 616, Greenbelt, Maryland 20771, USA
[3]Science Systems and Applications, Inc., Greenbelt, Maryland 20706, USA
*pwzhai@umbc.edu

Received 28 December 2022; revised 12 March 2023; accepted 28 March 2023; posted 28 March 2023; published 21 April 2023

We investigated the optimal number of independent parameters required to accurately represent spectral remote sensing reflectances (R_{rs}) by performing principal component analysis on quality controlled *in situ* and synthetic R_{rs} data. We found that retrieval algorithms should be able to retrieve no more than four free parameters from R_{rs} spectra for most ocean waters. In addition, we evaluated the performance of five different bio-optical models with different numbers of free parameters for the direct inversion of in-water inherent optical properties (IOPs) from *in situ* and synthetic R_{rs} data. The multi-parameter models showed similar performances regardless of the number of parameters. Considering the computational cost associated with larger parameter spaces, we recommend bio-optical models with three free parameters for the use of IOP or joint retrieval algorithms. © 2023 Optica Publishing Group

https://doi.org/10.1364/AO.484082

1. INTRODUCTION

Ocean color sensors, including the Sea-Viewing Wide Field of View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), Medium Resolution Imaging Spectrometer (MERIS), Visible Infrared Imaging Radiometer Suite (VIIRS), and Ocean and Land Color Instrument (OLCI), have provided a continuous and synoptic view of the global oceans for more than 20 years. Upcoming programs, such as the Plankton, Aerosol, Cloud, Ocean Ecosystem (PACE) mission, will further extend the ocean color data record with global, hyperspectral measurements of our home planet [1]. These observations have or will enhance our understanding of the ocean's ecological response to Earth's changing climate [2], its role in global carbon cycles [3], the dynamics of coastal water quality [4,5], the health of aquatic fisheries [6], and the distribution of aquatic primary production [7].

Ocean color remote sensing algorithms estimate spectral water-leaving radiances from the total measured radiances emanating from the coupled ocean and atmosphere system by performing atmospheric correction (AC), which removes atmospheric path radiance and ocean surface contributions from the total signal observed by the satellite instrument [8]. Water-leaving radiances are then used to derive aquatic inherent and apparent optical properties and biogeochemical properties. The performance of these ocean color retrievals depends on the accuracy of the AC process in the retrieval. Heritage AC algorithms [9] applied under the black pixel approximation [10] in the near infrared (NIR) work well for open ocean scenes, but are less reliable for scenes involving complex coastal waters, where the radiance in the NIR cannot be assumed zero or reliably estimated, or for scenes involving complex atmospheres, such as those that include absorbing aerosols [11–13]. In contrast, one-step, or joint retrieval, algorithms directly fit the sensor-observed radiance with forward model simulations, which are based on the radiative transfer theory in Coupled Atmosphere and Ocean Systems (CAOS). The least squares fitting algorithms are used to perturb the state parameters of the atmosphere and ocean properties to minimize the difference between measurements and model simulations [14,15]. Most operational algorithms applied to multispectral or hyperspectral radiance measurements from single-viewing ocean color instruments utilize a two-step approach, whereas one-step algorithms are a common approach for multispectral, multi-angle polarimeters (MAPs) [16,17]. Owing to the large information content in MAP measurements [18–20], one-step algorithms can be successful under conditions where two-step heritage algorithms fail [13] if the CAOS models can adequately represent the observed scene.

One important component of all ocean color remote sensing algorithms is the bio-optical model that represents and defines the composition of the water under observation. Typically, a bio-optical model is composed of analytical and empirical

expressions or tabulated values for inherent optical properties (IOPs; absorption and scattering coefficients; m^{-1}) of different hydrosols: phytoplankton particles, non-algal particles (NAPs), colored dissolved organic matter (CDOM), and pure sea water [21]. Waters where hydrosol IOPs and their magnitudes largely covary with the concentration of the photosynthetic pigment chlorophyll-a ([*Chla*]; $mg\,m^{-3}$) are collectively referred to as case I waters and typically represent clear, open ocean conditions. The remaining waters, whose IOPs do not correlate well with [*Chla*], are referred to as case II waters and require additional parameters beyond [*Chla*] to parametrize a bio-optical model [21,22].

Broadly speaking, empirical [23,24] and semi-analytical algorithm (SAA) [25–29] bio-optical models are generally used to estimate IOPs or biogeochemical quantities of interest from spectral remote sensing reflectances ($R_{rs}(\lambda)$; sr^{-1}), defined as water-leaving radiances scaled to the downwelling surface irradiance, with spectral dependence hereafter implied. Empirical approaches relate optical or biogeochemical variables statistically to R_{rs} based on observed or measured relationships. For example, [*Chla*] is commonly derived from a blue-to-green ratio of R_{rs} derived from a global dataset of *in situ* measurements [30]. In contrast, SAAs predefine spectral parametrizations of component IOPs and use spectral matching techniques to retrieve IOP magnitudes that best reconstruct the input R_{rs}, with the component spectral shapes often formulated via tabulated values or, for example, as power laws or exponentials with spectral slopes that are either taken as constants [25,31] or based on empirical R_{rs} relationships [26,29]. SAAs employ a combination of empirical and simplified solutions to the radiative transfer equation and are most often based on linear or nonlinear spectral optimization [25,31–33], step-wise spectral deconvolution [28,34], or bulk inversion [35].

In one-step AC retrieval algorithms, bio-optical models are used to simulate spectral ocean water reflectances and thus enable the separation of that signal from the atmospheric path radiance contribution in the AC process [36], or to retrieve in-water IOPs [15,37,38]. Many publications on one-step joint retrieval algorithms [15,36,37,39] assume a case I bio-optical model based on a single parameter: [*Chla*]. Only several algorithms currently adopt case II bio-optical models [14,40,41] with higher order parameterizations. In these studies, the one-step algorithms were shown to perform better over complex optical scenes when using case II bio-optical models in lieu of single parameter case I models [38].

The retrieval uncertainties of one-step joint retrieval algorithms partially depend on the size of the parameter space, including that of the bio-optical models [38]. The balance between the model fidelity and parameter space is crucial. In the multi-angular polarimetric ocean color (MAPOL) joint retrieval algorithm [14,38,42,43], a seven-parameter bio-optical model for coastal waters is implemented in the forward radiative transfer model [44,45]. If, however, the seven free parameters can be reduced to a smaller number, this would improve the stability as well as the efficiency of the retrieval algorithm, as the number of forward model evaluations increases with the number of retrieved parameters. It is theoretically important to understand how many retrievable parameters can be reliably and stably derived from ocean radiometric measurements.

The goal of this study is two-fold. The first part focuses on understanding the information content in R_{rs}, which can help to better constrain the parameter space of aquatic bio-optical models. The second part focuses on identifying an optimal bio-optical model(s) that may have the smallest possible set of parameters to result in greatest retrieval performance, defined here as a balance among retrieval accuracy, computational efficiency, and model robustness. The resultant bio-optical model(s) can be used in one-step joint retrieval algorithms [14,38] as well as those retrieving ocean IOPs from in-water spectral radiometric measurements [28,35,46]. Sathyendranath *et al.* [47] carried out a principal component analysis (PCA) based on simulated data to understand the information content in R_{rs}. The methodology is similar to our current work. However, in addition to the simulated dataset, our PCA also includes *in situ* measurements acquired from a large variety of global ocean waters (see Fig. 1), and thus we provide here a more comprehensive and representative assessment of the information content of R_{rs} and the potential number of free bio-optical model parameters that can be confidently retrieved. In the second part of our study, we provide a comprehensive analysis of R_{rs} retrieval performances from a set of bio-optical models with different numbers of free parameters, which is corroborated by a PCA. While previous studies have concluded that a three-parameter bio-optical model is generally sufficient for R_{rs} inversion [29,31], these studies did not fully explore the potential for retrieving additional information using higher order parameterizations. Wang *et al.* [48] showed that spectral shaped parameters are more difficult to retrieve relative to other bio-optical model parameters. Brewin *et al.* [49] and Werdell *et al.* [50] compared retrieval performances between two or more aquatic bio-optical models, but did not specifically relate the results with the number of parameters.

This paper includes two parts. In the first part, we attempt to identify the fewest number of independent parameters required to accurately represent spectral R_{rs}. A CPA was performed on four sets of *in situ* and synthetic R_{rs} data (Section 2) covering a large range of ocean waters. We evaluate both the explained variance of the principal components (PCs) and the effect of excluding the least significant PCs in reconstructing R_{rs}. In the second part, we study the performance of five different bio-optical models as applied in the direct inversion of R_{rs} to obtain the IOPs of ocean waters. Four of the models used in the study are case II models, and the remaining is a case I model based on a single parameter [*Chla*]. Our case II models differ in how they parametrize the particulate backscattering coefficient and parameter space. Each model has two versions under two different parameter spaces. The inversions are performed with both *in situ* and synthetic R_{rs} measurements (Section 2).

This paper is organized as follows. Section 2 overviews the *in situ* and synthetic datasets used in the study; Section 3 describes the methodology of the two studies, and PC analysis of R_{rs} data and R_{rs} inversions from five bio-optical models; Section 4 presents and discusses the results; and finally, Section 5 summarizes the conclusions.

2. DATA

In this study, we acquired *in situ* and simulated R_{rs}, the spectral absorption coefficients of NAP + CDOM (a_{dg}) and phytoplankton particles (a_{ph}), spectral backscattering coefficients of particulate matter (b_{bp}), and [*Chla*] data from (1) the NASA Bio-Optical Marine Algorithm Dataset (NOMAD) [51], (2) the Valente *et al.* [52] dataset (VAL), (3) the IOCCG Ocean Colour Algorithms Working Group dataset [21] (IOCCG), and (4) the simulated dataset by Craig *et al.* [53] (CRG). The NOMAD and VAL datasets include *in situ* measurements, while the IOCCG and CRG datasets were simulated using radiative transfer models defined by bio-optical relationships. These data cover a wide dynamic range of optical water types, from clear oceanic waters to highly turbid coastal waters. The spectral availability of *in situ* data is scarce. To maximize the sample size from *in situ* datasets we selected R_{rs} and IOP data at the most common five wavelengths, 410, 440, 510, 555, and 665 nm, from each dataset. In addition we selected R_{rs} and IOP data with eight wavelengths, 410, 440, 490, 510, 532, 555, 617, and 665 nm [54], to investigate sensitivities in the retrievability of ocean bio-optical parameters with increasing spectral information. The actual wavelengths obtained from each dataset may differ by ±5 nm. The global distribution of *in situ* R_{rs} datasets under five and eight wavelengths is shown in Fig. 1. The eight wavelength dataset obtained from NOMAD is geographically limited. All *in situ* R_{rs} data are filtered following a quality assurance (QA) system [55] to exclude any low quality R_{rs} with a QA score smaller than 0.5.

There are more than 4000 stations available in NOMAD. About 2300 R_{rs} spectra under five wavelengths and about 644 R_{rs} spectra under eight wavelengths remained after application of the QA test. VAL is a compilation of global bio-optical data from different sources including NOMAD, Aerosol Robotic Network—Ocean Color (AERONET-OC) [56], SeaWiFS Bio-optical Archive and Storage System (SeaBASS), etc., i.e., NOMAD is a subset of VAL. For a comparison with previous studies in the literature, we have presented NOMAD as a separate study. There are more than 50,000 cases in VAL at 611 hyperspectral wavelengths. Following the QA test, we obtained 12,167 and 3911 R_{rs} spectra under five and eight wavelengths, respectively. The sample sizes for IOP data from the two *in situ* datasets differ from those of R_{rs} due to the instrumental variabilities and availabilities in associated field campaigns.

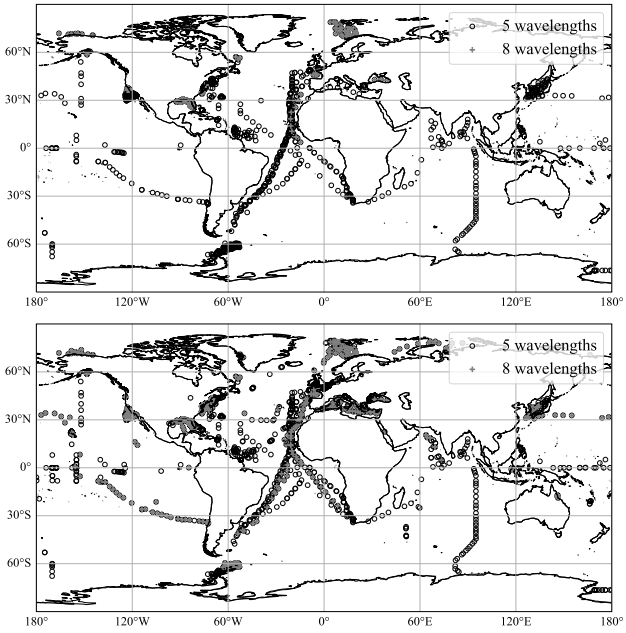

Fig. 1. Global geographic distribution of NOMAD (top) and VAL (bottom) *in situ* R_{rs} measurements. Circles: five wavelength dataset. Crosses: eight wavelength dataset.

The IOCCG synthetic data include 500 R_{rs} spectra generated using an array of IOP conditions defined by bio-optical relationships. We selected all available data to obtain datasets under five and eight wavelengths without following any QA tests. CRG is a hyperspectral synthetic dataset generated with IOP conditions based on phytoplankton absorption coefficient measurements from SeaBASS and eight other random parameters. It includes few R_{rs} spectra corresponding to inland waters, and we have excluded those in our study (based on NIR R_{rs} values). We selected 630 spectra for both five and eight wavelength datasets.

3. METHODOLOGY

A. Principal Component Analysis

In the first study, we estimated the optimal number of independent parameters that may be required to represent a given R_{rs} spectrum using the PCA technique. Initially, we performed PCA on *in situ* and synthetic R_{rs} datasets (Section 2) obtained at five and eight different wavelengths within the 410–665 nm range. R_{rs} is constructed using PCs in the form of

$$R_{rs}(\lambda) = R_{rs_0} + \sum_i \alpha_i P_i(\lambda), \tag{1}$$

where R_{rs_0} is the mean of R_{rs}, $P_i(\lambda)$ are the PCs, and α_i are their respective coefficients.

The method we followed to estimate the optimal number of independent parameters is as follows. For each dataset, we estimated a number of PCs similar to the number of wavelengths. Then we reconstructed each R_{rs} spectrum with the combinations of PCs according to their rank in the explained variance ratio. The first reconstruction is based on the PC capturing the highest variance, next is based on the combination of the first two PCs capturing highest variance, etc. We estimated the spectral residual by comparing the reproduced R_{rs} [Eq. (1)] with the corresponding true R_{rs}, which was then compared to the PACE R_{rs} uncertainty requirements [1] to quantify the fraction of reconstructed R_{rs} spectra with the residual at each wavelength within the PACE uncertainty requirement. The PACE R_{rs} uncertainty requirement is the community standard for allowable uncertainties, and we used it for convenience. It is spectrally dependent and originally defined for level 1 water-leaving reflectance ρ_w^+ (from 400 to 600 nm, the absolute uncertainty is 0.0020 and the relative uncertainty is 5.0%; from 600 to 710 nm, the absolute uncertainty is 0.00070 and the relative uncertainty is 10.0%). For R_{rs}, absolute uncertainty is obtained by scaling ρ_w^+ with π, ($R_{rs} = \rho_w^+/\pi$). If "relative uncertainty $\times R_{rs}$" for a particular wavelength was less than the absolute uncertainty, absolute uncertainty was used for the

comparison. Otherwise, the relative uncertainty was used for comparison.

The optimal number of free parameters is the sum of the optimal number of PCs and another parameter for R_{rs0}. To determine the impact of the number of wavelengths on the optimal number of free parameters, we compared the PCA results under five and eight wavelengths.

B. Performance Analysis of Bio-Optical Models

We adopted an SAA inversion approach for our bio-optical model framework, which derives [*Chla*] and IOPs from R_{rs}. The model parameters were estimated with spectral optimization. The SAA is described in the following.

1. Semi-Analytical Algorithm Expression

For optically deep waters, R_{rs} can be expressed as a function of total absorption $[a(\lambda)m^{-1}]$ and backscattering $[b_b(\lambda)m^{-1}]$ coefficients of the water column [57]:

$$R_{rs}(\lambda) \propto \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} = u(\lambda). \tag{2}$$

R_{rs} can be related to subsurface R_{rs} ($r_{rs}(\lambda, 0^-)$) by [34]

$$R_{rs}(\lambda) = \frac{0.52 r_{rs}(\lambda)}{1 - 1.7 r_{rs}(\lambda)}. \tag{3}$$

$r_{rs}(\lambda, 0^-)$ can be related to $u(\lambda)$ by [57]

$$r_{rs}(\lambda) = G_1 u(\lambda) + G_2 u(\lambda)^2. \tag{4}$$

We have adopted $G_1 = 0.0949$ and $G_2 = 0.0794$ [57]. Alternative G coefficients also exist in literature [34,58]. In general, $G_i(sr^{-1})$, $i = 1, 2$ is a spectrally dependent coefficient that varies with illumination condition and geometry, in-water IOPs (more specifically, particle phase functions), and sea surface properties such as wind speed and bidirectional effects [59]. The total absorption and backscattering coefficients can be expanded as the sum of absorption and backscattering coefficients of the independent subcomponents

$$a(\lambda) = a_w(\lambda) + a_{\mathrm{ph}}(\lambda) + a_g(\lambda) + a_d(\lambda) \tag{5}$$

and

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda), \tag{6}$$

where subscripts w, ph, g, and d refer to seawater, phytoplankton, CDOM, and NAP, respectively. In Eq. (6), we assume that CDOM has no contribution in scattering, and consider phytoplankton and NAP together as particulates (indicated in subscript p). Due to the similarity in spectral shapes, CDOM and NAP absorption coefficients (a_{dg}) are also typically combined as a single coefficient as a_{dg} in the remote sensing paradigm:

$$a(\lambda) = a_w(\lambda) + a_{\mathrm{ph}}(\lambda) + a_{dg}(\lambda). \tag{7}$$

Both $a_w(\lambda)$ [60] and $b_{bw}(\lambda)$ coefficients [61] are known from experimental measurements.

2. Bio-Optical Models

In this study, we explored the following five candidate bio-optical models: C2MP5, C2NP5, C2MP3, C2NP3, and C1P1 for optically deep waters. In the names, the first two characters indicate the water types, i.e., case II models (C2) and case I (C1) model. In case II models, the third letters M and N indicate different ways to represent the particulate backscattering coefficient. In M models, an analytical expression for b_{bp} is used, whereas in N models, the parametrization of the non-water extinction coefficient (c_{nw}) and backscattering fraction (B_p) are used. The last two characters in the model name show the number of free parameters, i.e., P3 stands for three free parameters. We have four case II models, and C1P1 is a single parameter model based on [*Chla*] only. A detailed description of the bio-optical models is given below.

C2MP5 [Eqs. (8)–(10)] is a five-parameter case II bio-optical model adopted from [14]:

$$a_{\mathrm{ph}}(\lambda) = A_{\mathrm{ph}}[Chla]^{E_{\mathrm{ph}}(\lambda)}, \tag{8}$$

$$a_{dg}(\lambda) = a_{dg}(440)\exp[-S_{dg}(\lambda - 440)], \tag{9}$$

$$b_{bp}(\lambda) = b_{bp}(660)\left(\frac{\lambda}{660}\right)^{-S_{bp}}, \tag{10}$$

where $a_{\mathrm{ph}}(\lambda)$ is parameterized in terms of [*Chla*] using A_{ph} and E_{ph} coefficients obtained from [62]; $a_{dg}(440)$ is the absorption coefficient of CDOM + NAP at 440 nm; $b_{bp}(660)$ is the backscattering coefficient of ocean particles at 660 nm; S_{dg} is the spectral exponential slope of a_{dg} in nm^{-1}, and S_{bp} is the spectral slope of the power law function of b_{bp}. The five free parameters are [*Chla*], $a_{dg}(440)$, $b_{bp}(660)$, S_{dg}, and S_{bp}.

The C2NP5 model [Eqs. (11)–(13)] uses a power law function to express c_{nw} [63] instead of b_{bp}:

$$c_{nw}(\lambda) = c_{nw}(660)\left(\frac{\lambda}{660}\right)^{-\gamma}, \tag{11}$$

$$b_p(\lambda) = c_{nw}(\lambda) - (a_{\mathrm{ph}}(\lambda) + a_{dg}(\lambda)), \tag{12}$$

$$b_{bp}(\lambda) = B_p \times b_p(\lambda), \tag{13}$$

where $c_{nw}(660)$; m^{-1} is the attenuation coefficient of non-water components at 660 nm, and γ is the spectral slope of c_{nw}. We have assumed spectrally invariant backscattering fraction B_p for all oceanic waters [64]. The rest of the model is similar to C2MP5, except we have fixed S_{dg} at 0.018 nm^{-1} [Eq. (9)] [65]. The five free parameters are [*Chla*], $a_{dg}(440)$, $c_{nw}(660)$, γ, and B_p. The rationale for this approach relative to C2MP5 arises from the greater likelihood that $c_{nw}(\lambda)$ follows a power law than $b_{bp}(\lambda)$ [66].

C2MP3 is the three-parameter model derived from the C2MP5 model [Eqs. (8)–(10)]. We fixed the spectral slopes S_{dg} at 0.018 nm^{-1} and S_{bp} at 0.3. The value of S_{bp} was obtained from a sensitivity analysis carried out with NOMAD and IOCCG data. In the sensitivity analysis, we performed the R_{rs} inversion with 30 S_{bp} values within 0–2.5 range, and found the

C2MP5	C2NP5
$a_{ph}(\lambda) = A_{ph}(\lambda)[\boldsymbol{Chla}]^{E_{ph}(\lambda)}$ $a_{dg}(\lambda) = \boldsymbol{a_{dg}(440)}\exp[-\boldsymbol{S_{dg}}(\lambda - 440)]$ $b_{bp}(\lambda) = \boldsymbol{b_{bp}(660)}\left(\frac{\lambda}{660}\right)^{-\boldsymbol{S_{bp}}}$	$a_{ph}(\lambda) = A_{ph}(\lambda)[\boldsymbol{Chla}]^{E_{ph}(\lambda)}$ $a_{dg}(\lambda) = \boldsymbol{a_{dg}(440)}\exp[-0.018(\lambda - 440)]$ $c_{nw}(\lambda) = \boldsymbol{c_{nw}(660)}\left(\frac{\lambda}{660}\right)^{-\gamma}$ $\boldsymbol{B_p}$
C2MP3	**C2NP3**
$a_{ph}(\lambda) = A_{ph}(\lambda)[\boldsymbol{Chla}]^{E_{ph}(\lambda)}$ $a_{dg}(\lambda) = \boldsymbol{a_{dg}(440)}\exp[-0.018(\lambda - 440)]$ $b_{bp}(\lambda) = \boldsymbol{b_{bp}(660)}\left(\frac{\lambda}{660}\right)^{-0.3}$	$a_{ph}(\lambda) = A_{ph}(\lambda)[\boldsymbol{Chla}]^{E_{ph}(\lambda)}$ $a_{dg}(\lambda) = \boldsymbol{a_{dg}(440)}\exp[-0.018(\lambda - 440)]$ $c_{nw}(\lambda) = \boldsymbol{c_{nw}(660)}\left(\frac{\lambda}{660}\right)^{-0.8}$ $B_p = 0.01$
C1P1	
[**Chla**] based model	

Fig. 2. Summary of five different bio-optical models. Free parameters of each model are indicated in bold.

use of $S_{bp} = 0.3$ leads to the best retrieval accuracy. However, in literature [31,32], 1.0 is more commonly used for S_{bp}. The difference in the value we obtained might be due to the geolocation distribution of NOMAD data, as S_{bp} tends to approach zero for coastal waters. The free parameters of the model are $[Chla]$, $a_{dg}(440)$, and $b_{bp}(660)$.

C2NP3 is the three-parameter model derived from the C2NP5 model [Eqs. (8), (9), and (11)]. We fixed B_p at 0.01 [Eq. (13)] [64] and γ at 0.8 [Eq. (11)]. The value of γ was obtained by a similar procedure as used in determining the value of S_{bp}. The free parameters are $[Chla]$, $a_{dg}(440)$, and $c_{nw}(660)$.

C1P1 is a $[Chla]$ based single parameter case I model [38], which is not representative of the complex case II waters. The expressions for component IOPs in this model are derived from the C2MP5 model [Eqs. (8)–(10)]. The absorption coefficient of phytoplankton a_{ph} is the same as in Eq. (8). We adopted the same expressions for a_{dg} as in the C2MP5 model, but the parameters are specified in terms of $[Chla]$ due to the sole phytoplankton dependence. S_{dg} is also fixed at 0.018 nm^{-1}:

$$a_{dg}(440) = p_2 a_{\mathrm{ph}}(440), \tag{14}$$

$$p_2 = 0.3 + \frac{(5.7 \times 0.5 a_{\mathrm{ph}}(440))}{(0.02 + a_{\mathrm{ph}}(440)}. \tag{15}$$

Similarly, b_{bp} is also contributed only from phytoplankton and is expressed in terms of $[Chla]$ [67]:

$$b_{bp}(\lambda) = b_p(\lambda) \times B_p, \tag{16}$$

$$b_p(\lambda) = b_p(660)(\lambda/(660)^{-S_{bp}}, \tag{17}$$

$$b_p(660) = 0.347[Chla]^{0.766}. \tag{18}$$

For $0.02 < [Chla] < 2$ mgm^{-3},

$$S_{bp} = -0.5(\log_{10}[Chla] - 0.3). \tag{19}$$

For $[Chla] > 2$ mgm^{-3}, $S_{bp} = 0$,

$$B_p = 0.002 + 0.01(0.5 - 0.25\ \log_{10}[Chla]). \tag{20}$$

Figure 2 shows the summary of five bio-optical models.

Table 1. Upper and Lower Boundaries of Parameters from Each Model

Parameter	Model	Lower/Upper Boundaries
$[Chla](\mathrm{mgm}^{-3})$	All models	0.0, 30.0
$a_{dg}(440)(m^{-1})$	C2MP5, C2NP5, C2MP3, C2NP3	0.0, 2.5
$S_{dg}(nm^{-1})$	C2MP5	0.005, 0.02
$b_{bp}(660)(m^{-1})$	C2MP5, C2MP3	0.0, 0.1
S_{bp}	C2MP5	0.0, 2.5
$c_{nw}(660)(m^{-1})$	C2NP5, C2NP3	0.0, 3.0
γ	C2NP5	0.0, 2.5
B_p	C2NP5	0.005, 0.06

3. R_{rs} Inversion

We retrieved the unknown free parameters from each model through Levenburg Marquardt (LM) non-linear least squares optimization [68]. LM does not explicitly support bounds on parameters. We used an LM algorithm capable of handling parameter bounds in terms of periodic boundary conditions [69]. The upper and lower boundaries of each parameter are listed in Table 1. The boundary values were derived from data given in Section 2 and previous literature [28,34,41,64,70]. The objective function [Eq. (21)] χ^2 of optimization is

$$\chi^2 = \sum_{i=1}^{N_\lambda} \frac{[\overline{R_{rs}}(\lambda_i) - R_{rs}(\lambda_i)]^2}{\sigma(\lambda_i)^2}, \tag{21}$$

where $\overline{R_{rs}}$ is the remote sensing reflectance calculated from SAA [Eqs. (1)–(3)] based on input parameters from each bio-optical model, R_{rs} is the data, and N_λ is the number of wavelengths. Due to the unavailability of R_{rs} uncertainties ($\sigma(\lambda_i)$; sr^{-1}), we carried out an unweighted optimization ($\sigma(\lambda_i) = 1$). The maximum number of function evaluations was 1000. The convergence was reached if the relative error in the desired sum of squares was less than 10^{-7}. We performed R_{rs} inversions using both five and eight wavelength R_{rs} data under each dataset. The validity of the retrievals was based on the criterion from [29]:

1. $-0.05 b_{bw}(\lambda) \le b_{bp}(\lambda) \le 0.1 m^{-1}$,
2. $-0.05 a_w(\lambda) \le a_{dg}(\lambda) \le 5 m^{-1}$,
3. $-0.05 a_w(\lambda) \le a_\phi(\lambda) \le 5 m^{-1}$,
4. $\Delta R_{rs} \le 33\%$,

where

$$\Delta R_{rs} = 100\% \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} \frac{\left[\overline{R_{rs}}(\lambda_i) - R_{rs}(\lambda_i)\right]}{R_{rs}(\lambda_i)}. \tag{22}$$

As the R_{rs} values in longer wavelengths are typically smaller, we considered only wavelengths ranging from 400 to 600 nm to calculate ΔR_{rs}. Considering the possible measurement uncertainties in *in situ* R_{rs} data and modeling uncertainties in the synthetic R_{rs} data, we considered negative IOP retrievals within the aforementioned limitations to be valid. Cases resulting in negative $\overline{R_{rs}}$ or $\overline{R_{rs}}$ that never reached convergence were regarded as invalid.

4. Validation Analysis

We performed a statistical evaluation to analyze the performances of each bio-optical model to accurately invert the input R_{rs} spectra. The performance is evaluated based on retrieved IOPs. We adopted type II linear regression statistics [71] and Bland and Altman (BA) analysis [72] to evaluate the accuracy of retrieved IOPs against true IOPs [$a_{dg}(440)$, $b_{bp}(660)$, and $a_{\mathrm{ph}}(440)$]. Since the IOP data are log normally distributed, the analyses are based on log transformed data. The correlation is based on Spearman's rank correlation coefficient. (Figure 6 provides common statistical measures in the ocean color community (correlation coefficient, slope, and unbiased RMSE), and Fig. 7 provides BA analysis [limits of agreement (LOA) and mean bias]).

4. RESULTS AND DISCUSSION

A. Principal Component Analysis

We performed PCA on R_{rs} datasets obtained at five and eight different wavelengths within the visible region of the electromagnetic spectrum. For all datasets, the first three PCs capture more than 95% of the variance in R_{rs} data (Fig. 3). In comparison to the PCA from five wavelength R_{rs} data, the first PC from eight wavelength R_{rs} data captures more variance. The higher order PCs are slightly significant under eight wavelength R_{rs} data compared to five wavelengths, due to the different wavelength selection.

Using Eq. (1), the PCs can be used to reconstruct the R_{rs} spectrum. Figure 4 shows the fraction of reconstructed R_{rs} spectra whose residual is within the PACE baseline R_{rs} uncertainty requirement. The residual is defined as the difference between true R_{rs} and reconstructed R_{rs} from different numbers of PCs. It shows that the numbers of PCs required to construct R_{rs} spectra are similar across the four datasets.

The results under the five wavelength datasets indicate that three PCs can be used to accurately reconstruct R_{rs}. For wavelengths shorter than 600 nm, around 95% of the four datasets meet the PACE baseline uncertainty requirement. For the long wavelength at 665 nm, three PCs can explain around 75% and 60% of the variability for *in situ* and synthetic datasets, respectively. With eight wavelengths, around 90% of the four datasets can be explained by three PCs when the wavelengths are shorter than 600 nm. For longer wavelengths, it drops to around 80%. The results show that the minimum number of PCs required for optimal representation of R_{rs} ($N_{\mathrm{PC}} = 3$) is largely unaffected by the number of visible wavelengths used in PCA. Overall, this translates to about four independent parameters for optimal representation of spectral R_{rs} [Eq. (1)].

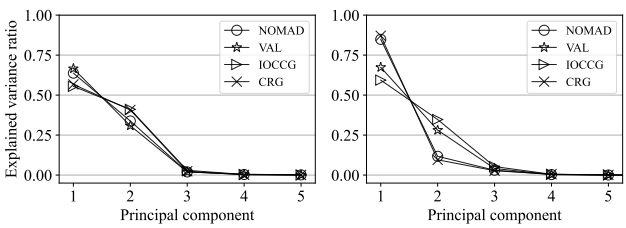

Fig. 3. Explained variance ratio of each PC obtained from R_{rs} data at five (left) and eight wavelength (right) datasets.

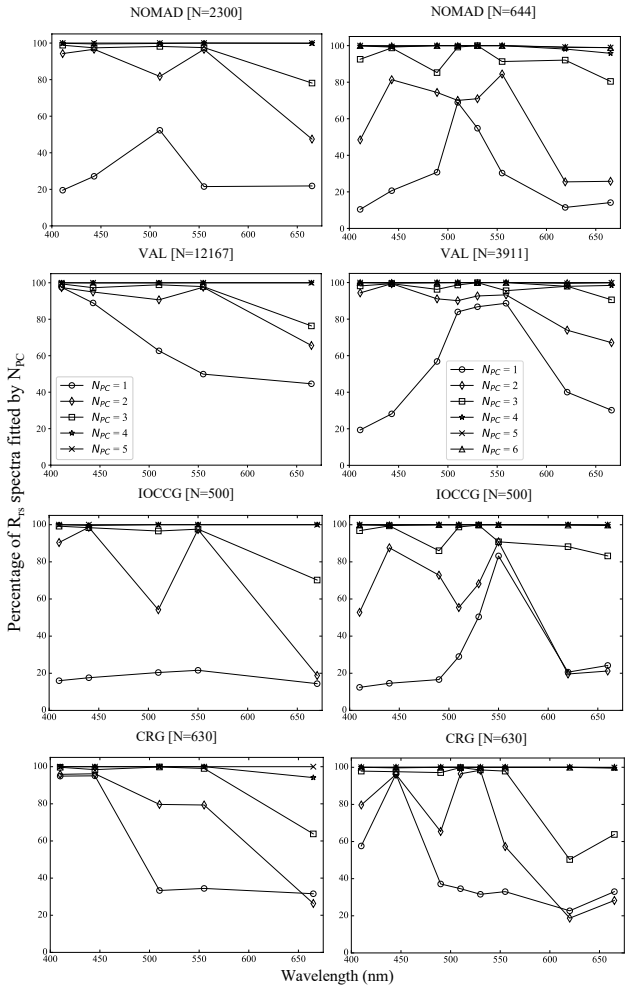

Fig. 4. Percentage of R_{rs} spectra (left: five wavelengths; right: eight wavelengths) whose residual level is smaller than the PACE baseline uncertainty requirement. The residual is defined as the difference between true R_{rs} and reconstructed R_{rs} from different numbers of PCs. PCs are ranked with respect to their explained variance (1, PC with the highest variance; 2, PC combination of first and second highest variances, etc.). N_{PC} is the number of PCs used in the reconstruction.

Hyperspectral R_{rs} data capture more spectral features such as the contribution from chlorophyll-a fluorescence (around 685 nm). These features are less evident in synthetic data. The availability of high quality *in situ* hyperspectral R_{rs} measurements is extremely limited, so we have not examined the impact of hyperspectral wavelengths on PCA. The PCA study based on *in situ* hyperspectral R_{rs} measurements [73] (sample size = 191) shows similar degrees of freedom (DoFs) for both hyperspectral and multi-spectral R_{rs} (using hyperspectral R_{rs} coarsened for MODIS wave bands), which agrees with our finding (i.e., four independent parameters).

Since PCs have defined their own space, they are less interpretable and cannot be related to physical variables. Furthermore, the parameters we adopt in bio-optical models are often correlated with each other. Hence, the optimal number of independent parameters may not be directly related to the number of optimal model parameters we use in a bio-optical model for retrieval. The significance of the PCA is two-fold. On one hand, it can be used as a guideline for the maximum number of

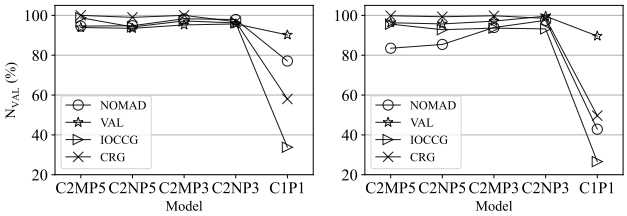

Fig. 5. Percentage of valid retrievals, N_{VAL}(%), under each model with the four datasets: NOMAD, VAL, IOCCG, and CDG. Left: five wavelength datasets. Right: eight wavelength datasets.

free retrieval parameters. This will provide crucial information on designing aquatic bio-optical models for hyperspectral R_{rs} spectra. In addition, PCs can be used as an alternate method to replace bio-optical models [74].

B. Performance Analysis of Bio-Optical Models in Inversion

We used the five bio-optical models (C1P1, C2NP3, C2MP3, C2NP5, and C2MP5) in retrieving IOPs with our inversion algorithm as summarized in Section 3.B. The validation statistics from the inversion under the four datasets are summarized in Figs. 5 and 6. The statistics are based on retrieved and true IOPs: $a_{dg}(440)$, $a_{\mathrm{ph}}(440)$, and $b_{bp}(660)$ from the four datasets: NOMAD, VAL, IOCCG, and CRG using only successful retrievals (Fig. 5).

Overall, the R_{rs} inversion results from case II bio-optical models (Fig. 5) show performance superior to that of the case I model, confirming that the C1P1 [*Chla*] based model is less applicable over a large range of waters. Hereafter, the results from C1P1 will not be included in the discussion. More than 90% of retrievals from case II models were valid for the five and eight wavelength datasets. An exception was seen in the eight wavelength NOMAD dataset, which is geographically limited. This may affect the reported inversion performance. For IOCCG and CRG, more than 95% of cases were valid, and for NOMAD and VAL, it was more than 90%. None of the inversions from any datasets resulted in convergence failures. For both five and eight wavelength datasets, an average of 3% and 0.5% cases failed when the IOPs were invalid and when $\Delta R_{rs} > 33\%$, respectively. Based on the validity, inversions from P3 models correspond to less than 2% invalid cases. The C2NP5 and C2NP3 models resulted in few invalid cases, which is attributed to negative R_{rs}, originated from the retrieval of negative b_{bp} values, as we consider the difference between c_{nw} and $a_{\mathrm{ph}} + a_{dg}$ to derive b_{bp}.

For all case II models, inversion results are similar under the five and eight wavelength datasets. Inversions based on the eight wavelength datasets could be less accurate compared to the five wavelength inversions due to the increased tendency of overfitting. Note that the inversion results from the NOMAD eight wavelength dataset are expected to be different from the five wavelength dataset due to the limited geographical distribution. The variability in statistics (Fig. 6) is more prominent for *in situ* data compared to synthetic data, specifically in $a_{dg}(440)$ and $a_{dg}(440)$. The lowest variability is seen for $b_{bp}(660)$. The *in situ* measurements include measurement uncertainties. On the other hand synthetic measurements are generated with

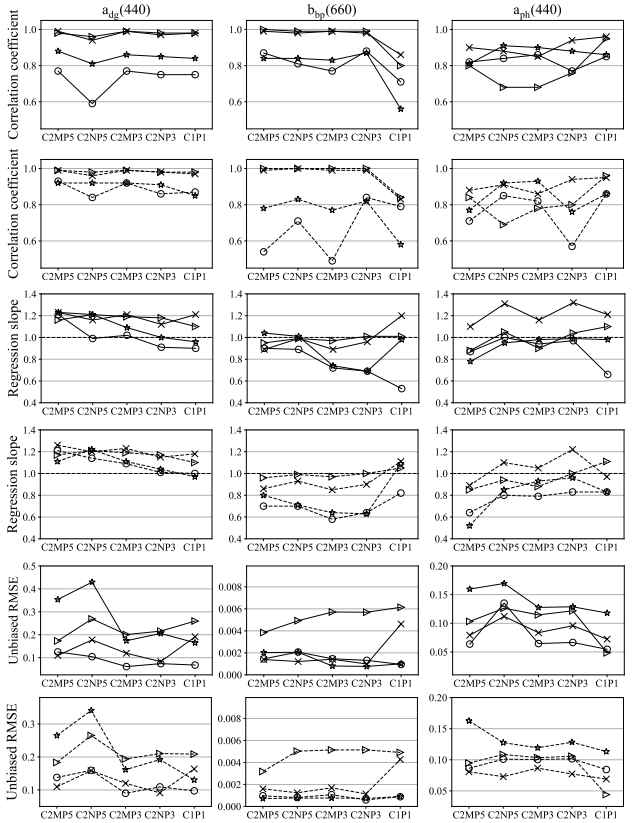

Fig. 6. Regression statistics (correlation coefficient, regression slope, and unbiased RMSE) from the five bio-optical models using NOMAD, VAL, IOCCG, and CDG datasets. Unbiased RMSE is the root mean square error corrected for mean bias. Data are log transformed except for unbiased RMSE. Correlation coefficient is the Spearman rank correlation coefficient. Different markers indicate each dataset. (O: NOMAD, ⋆: VAL, ▷: IOCCG, and ×: CRG) Solid lines: statistics from five wavelength datasets. Dashed lines: statistics from eight wavelength datasets.

parametrizations similar to what we have employed in our bio-optical models. For these reasons, we expect the synthetic data to show more robust results compared to *in situ* data. Out of the four models, C2MP5 shows the highest correlation with true IOPs ($a_{dg}(440)$). C2MP3 and C2NP3 are comparable. Both C2MP3 and C2NP3 show regression slopes closer to one [$a_{dg}(440)$ and $a_{\mathrm{ph}}(440)$]. The lowest unbiased RMSE is obtained for C2MP3 and C2NP3. However, overall statistics suggest that the P3 case II models are comparable to P5 case II models.

Referring to literature, the inversion statistics from NOMAD and IOCCG data under the case II models are comparable to the generalized IOP model—default configuration (GIOP-DC) [29], but with slight differences in the regression slope. This could be explained by the R_{rs} dynamic empirical spectral slope relationship for b_{bp} adopted in GIOP-DC, which is based on an R_{rs} band ratio [34]. Dynamic spectral slopes can better represent natural waters, but cannot be directly used in one-step joint retrieval algorithms where the input is not R_{rs}.

For further validation, we performed BA analysis on retrieved and true IOPs, considering only succssful retrievals (Fig. 5). Given that NOMAD data are widely used and consist of high

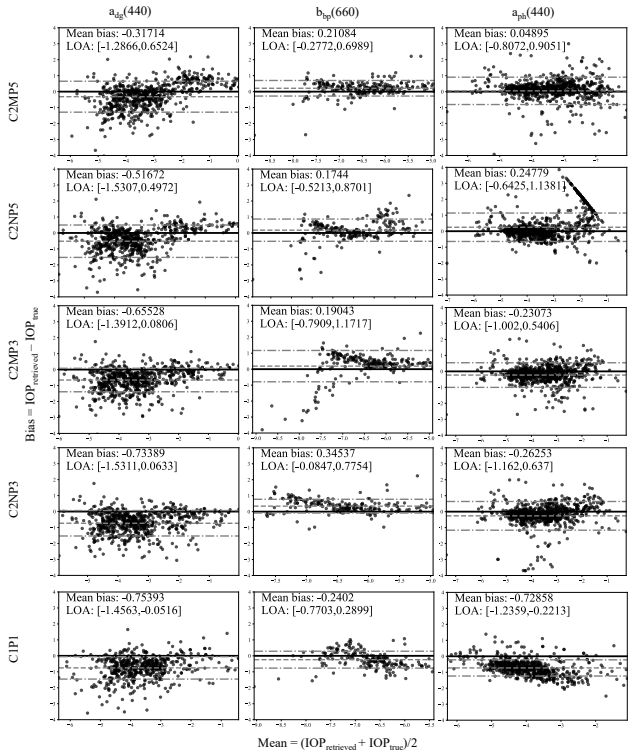

Fig. 7. Differences versus mean plot of log transformed IOP_{true} and $IOP_{retrieved}$. The dashed line shows the mean bias, and the dotted-dashed lines indicate the LOA. Bias is represented by the gap between mean bias line and line of equality.

quality *in situ* measurements, we narrowed the BA analysis to the NOMAD five wavelength dataset. The LOA and mean bias were calculated based on the bias ($IOP_{retrieved} - IOP_{true}$) and mean ($(IOP_{retrieved} + IOP_{true})/2$) from log transformed data. Figure 7 shows the results from BA analysis. We saw slightly reduced mean biases in the P5 models compared to the P3 models. It is more prominent for $a_{dg}(440)$, but the LOA are smaller for $a_{dg}(440)$ with P3 models. In the two P3 models, C2MP3 shows smaller overall mean biases than C2NP3. The C1P1 model shows a similar performance in retrieving $a_{dg}(440)$ as case II models. However, its performance is much worse than case II models in terms of $a_{ph}(440)$ and $b_{bp}(660)$; thus it is not recommended in IOP retrieval algorithms. There is a light trend in the retrieval of a_{dg} (C2MP5) and a_{ph} (C1P1), which may be due to the representativeness of the bio-optical models used. It remains a future research subject to fine tune the retrieval algorithms to reduce the retrieval bias and uncertainties.

All the analyses based on the five and/or eight wavelength datasets show that the R_{rs} inversion performances across the four case II models are similar regardless of the number of parameters (P3 or P5). This suggests that there is a possibility to reduce the size of bio-optical model parameter space of one-step retrieval algorithms while preserving the retrieval performance in terms of uncertainty, and improving efficiency.

C. Limitations of the Study

We performed this study in search of an optimal bio-optical model for retrieving aquatic IOPs from spectral R_{rs} with a particular focus on their use in one-step AC algorithms. The optimal bio-optical model for such use in one-step AC should result in the greatest inversion performance while having the smallest possible number of free parameters. Due to the unavailability of R_{rs} measurement uncertainties, we carried out unweighted R_{rs} inversions. This might have affected the model performances discussed above. In addition, modeling errors can arise due to the spectral shape assumptions of a_{dg}, b_{bp}, and c_{nw}, assumption of spectrally invariant backscattering fraction, fixed spectral slopes, and assumptions of R_{rs}–IOP relationships [e.g., the G_1 and G_2 values in Eq. (4)]. We have adopted the model of Bricaud *et al.* to represent a_{ph} spectra, which is an average of *in situ* measurements from ocean waters. Hence it may not necessarily account for all types of waters including in-land waters. Our study does not consider in-land waters. We acknowledge that none of these models ideally represents all types of water conditions. Finally, the optimization techniques adopted in the inversions may also cause some uncertainty.

5. CONCLUSIONS

We have performed two studies to understand the implication of bio-optical model parameter space in the retrieval of aquatic IOPs. The study is two-fold. In the first part, we examined the optimal number of independent parameters that may be required to represent the spectral behavior of R_{rs} using PCA. The study was based on the four datasets NOMAD, VAL, IOCCG, and CRG under two different selections of wavelengths. It was found that a minimum of three PCs is required to accurately represent an R_{rs} spectrum at either five or eight wavelengths within 410–665 nm range, regardless of the water condition. This suggests that the maximum number of retrievable free parameters is approximately four for most ocean waters. Due to the possible correlations within model parameters, the minimum number of model parameters required to represent R_{rs} is likely higher than the minimum number of free and independent parameters. This study also demonstrates the use of PCA as a tool to evaluate the free parameter selection of bio-optical models.

In the second part of the study, we evaluated the performances of five different bio-optical models (four case II water models, C2MP5, C2NP5, C2MP3, and C2NP3, and one case I water model, C1P1) in retrieving aquatic IOPs from R_{rs} data at five and eight wavelengths. Regardless of the number of parameters and the number of wavelengths used in the inversion, all case II models (three and five parameters) showed a similar performance in R_{rs} inversion with SAA. This implied the possibility to reduce bio-optical model parameter space without substantially altering the retrieval performances. In the two three-parameter models, C2MP3 shows smaller mean biases and range of variability than C2NP3.

In the future, we intend to implement a three-parameter bio-optical model in the MAPOL algorithm [14,38]. MAPOL is a joint retrieval algorithm that retrieves aerosol and ocean properties from MAP measurements. The algorithm employs ocean color bio-optical models to characterize the water-leaving radiance. The number of bio-optical model parameters directly affects the accuracy of AC based on MAPOL. As a validation study, we will perform polarimetric retrievals with the MAP

measurements from Research Scanning Polarimeter (RSP) [75,76], AirHARP [77], and SPEX airborne [78,79] to demonstrate C2MP3's capability in one-step retrieval algorithms. This study will further facilitate the ocean color remote sensing community with necessary guidance to take future directions in advancing bio-optical models for one-step algorithms and SAAs.

Funding. National Aeronautics and Space Administration (80NSSC20M0227); National Science Foundation (CNS-0821258, CNS-1228778, OAC-1726023, DMS-0821311).

Acknowledgment. The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program and the SCREMS program, with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Refs. [21,51,52,53].

REFERENCES

1. P. J. Werdell, M. J. Behrenfeld, P. S. Bontempi, E. Boss, B. Cairns, G. T. Davis, B. A. Franz, U. B. Gliese, E. T. Gorman, O. Hasekamp, K. D. Knobelspiesse, A. Mannino, J. V. Martins, C. R. McClain, G. Meister, and L. A. Remer, "The Plankton, Aerosol, Cloud, Ocean Ecosystem Mission: status, science, advances," Bull. Am. Meteorol. Soc. **100**, 1775–1794 (2019).
2. H. Gitay, S. Brown, W. Easterlin, and B. Jallow, "5. Ecosystems and their goods and services," in *Intergovernmental Panel on Climate Change 2001: Impacts, Adaptation and Vulnerability. Working Group II of the Intergovernmental Panel on Climate Change* (Cambridge University, 2001), pp. 238–342.
3. S. Levitus, J. I. Antonov, T. P. Boyer, and C. Stephens, "Warming of the world ocean," Science **287**, 2225–2229 (2000).
4. V. Brando and A. Dekker, "Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality," IEEE Trans. Geosci. Remote Sens. **41**, 1378–1387 (2003).
5. D. Blondeau-Patissier, J. F. Gower, A. G. Dekker, S. R. Phinn, and V. E. Brando, "A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans," Prog. Oceanogr. **123**, 123–144 (2014).
6. K. R. Stuart and M. Drawbridge, "The effect of light intensity and green water on survival and growth of cultured larval California yellowtail (*Seriola lalandi*)," Aquaculture **321**, 152–156 (2011).
7. T. Platt, S. Sathyendranath, M.-H. Forget, G. White, C. Caverhill, H. Bouman, E. Devred, and S. Son, "Diagnostic properties of phytoplankton time series from remote sensing," Remote Sens. Environ. **112**, 428–439 (2008).
8. C. Mobley, J. Werdell, B. Franz, Z. Ahmad, and S. Bailey, "Atmospheric correction for satellite ocean color radiometry," Technical Report NASA/TM–2016-217551 (National Aeronautics and Space Administration, 2016).
9. H. R. Gordon and M. Wang, "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm," Appl. Opt. **33**, 443–452 (1994).
10. S. W. Bailey, B. A. Franz, and P. J. Werdell, "Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing," Opt. Express **18**, 7521–7527 (2010).
11. IOCCG, "Remote sensing of ocean colour in coastal, and other optically-complex, waters," No. 3 of Reports of the International Ocean Colour Coordinating Group (IOCCG, 2000).
12. IOCCG, "Atmospheric correction for remotely-sensed ocean-colour products," No. 10 of Reports of the International Ocean Colour Coordinating Group (IOCCG, 2010).
13. R. J. Frouin, B. A. Franz, A. Ibrahim, *et al.*, "Atmospheric correction of satellite ocean-color imagery during the PACE era," Front. Earth Sci. **7**, 145 (2019).
14. M. Gao, P.-W. Zhai, B. Franz, Y. Hu, K. Knobelspiesse, P. J. Werdell, A. Ibrahim, F. Xu, and B. Cairns, "Retrieval of aerosol properties and water-leaving reflectance from multi-angular polarimetric measurements over coastal waters," Opt. Express **26**, 8968–8989 (2018).
15. S. Stamnes, C. Hostetler, R. Ferrare, S. Burton, X. Liu, J. Hair, Y. Hu, A. Wasilewski, W. Martin, B. Van Diedenhoven, J. Chowdhary, I. Cetinić, L. K. Berg, K. Stamnes, and B. Cairns, "Simultaneous polarimeter retrievals of microphysical aerosol and ocean color parameters from the "MAPP" algorithm with comparison to high-spectral-resolution lidar aerosol and ocean products," Appl. Opt. **57**, 2394–2413 (2018).
16. A. A. Kokhanovsky, "The modern aerosol retrieval algorithms based on the simultaneous measurements of the intensity and polarization of reflected solar light: a review," Front. Environ. Sci. **3**, 4 (2015).
17. O. Dubovik, Z. Li, M. I. Mishchenko, *et al.*, "Polarimetric remote sensing of atmospheric aerosols: instruments, methodologies, results, and perspectives," J. Quant. Spectrosc. Radiat. Transfer **224**, 474–511 (2019).
18. P.-Y. Deschamps, F.-M. Bréon, M. Leroy, A. Podaire, A. Bricaud, J.-C. Buriez, and G. Seze, "The POLDER mission: instrument characteristics and scientific objectives," IEEE Trans. Geosci. Remote Sens. **32**, 598–615 (1994).
19. D. J. Diner, J. C. Beckert, T. H. Reilly, C. J. Bruegge, J. E. Conel, R. A. Kahn, J. V. Martonchik, T. P. Ackerman, R. Davies, S. A. Gerstl, H. R. Gordon, J.-P. Muller, R. B. Myneni, P. J. Sellers, B. Pinty, and M. M. Verstraete, "Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview," IEEE Trans. Geosci. Remote Sens. **36**, 1072–1087 (1998).
20. K. Knobelspiesse, B. Cairns, M. Mishchenko, J. Chowdhary, K. Tsigaridis, B. van Diedenhoven, W. Martin, M. Ottaviani, and M. Alexandrov, "Analysis of fine-mode aerosol retrieval capabilities by different passive remote sensing instrument designs," Opt. Express **20**, 21457–21484 (2012).
21. IOCCG, "Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and application," No. 5 of Reports of the International Ocean Colour Coordinating Group (IOCCG, 2006).
22. P. J. Werdell, L. I. McKinna, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, D. Stramski, J. M. Sullivan, M. S. Twardowski, M. Tzortziou, and X. Zhang, "An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing," Prog. Oceanogr. **160**, 186–212 (2018).
23. J. E. O'Reilly, S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. McClain, "Ocean color chlorophyll algorithms for SeaWiFS," J. Geophys. Res. Oceans **103**, 24937–24953 (1998).
24. C. Hu, Z. Lee, and B. Franz, "Chlorophyll algorithms for oligotrophic oceans: a novel approach based on three-band reflectance difference," J. Geophys. Res. Oceans **117** (2012).
25. S. A. Garver and D. A. Siegel, "Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea," J. Geophys. Res. Oceans **102**, 18607–18625 (1997).
26. K. L. Carder, F. Chen, Z. Lee, S. Hawes, and D. Kamykowski, "Semianalytic moderate-resolution imaging spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures," J. Geophys. Res. Oceans **104**, 5403–5421 (1999).
27. C. S. Roesler and E. Boss, "Spectral beam attenuation coefficient retrieved from ocean color inversion," Geophys. Res. Lett. **30**, 1468 (2003).
28. T. J. Smyth, G. F. Moore, T. Hirata, and J. Aiken, "Semianalytical model for the derivation of ocean color inherent optical properties: description, implementation, and performance assessment," Appl. Opt. **45**, 8116–8131 (2006).
29. P. J. Werdell, B. A. Franz, S. W. Bailey, *et al.*, "Generalized ocean color inversion model for retrieving marine inherent optical properties," Appl. Opt. **52**, 2019–2037 (2013).

30. J. E. O'Reilly and P. J. Werdell, "Chlorophyll algorithms for ocean color sensors—OC4, OC5 & OC6," Remote Sens. Environ. **229**, 32–47 (2019).
31. S. Maritorena, D. A. Siegel, and A. R. Peterson, "Optimization of a semianalytical ocean color model for global-scale applications," Appl. Opt. **41**, 2705–2714 (2002).
32. C. S. Roesler and M. J. Perry, "In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance," J. Geophys. Res. Oceans **100**, 13279–13294 (1995).
33. F. E. Hoge and P. E. Lyon, "Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: an analysis of model and radiance measurement errors," J. Geophys. Res. Oceans **101**, 16631–16648 (1996).
34. Z. Lee, K. L. Carder, and R. A. Arnone, "Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters," Appl. Opt. **41**, 5755–5772 (2002).
35. H. Loisel and D. Stramski, "Estimation of the inherent optical properties of natural waters from the irradiance attenuation coefficient and reflectance in the presence of Raman scattering," Appl. Opt. **39**, 3001–3011 (2000).
36. F. Xu, O. Dubovik, P.-W. Zhai, D. J. Diner, O. V. Kalashnikova, F. C. Seidel, P. Litvinov, A. Bovchaliuk, M. J. Garay, G. van Harten, and A. B. Davis, "Joint retrieval of aerosol and water-leaving radiance from multispectral, multiangular and polarimetric measurements over ocean," Atmos. Meas. Tech. **9**, 2877–2907 (2016).
37. J. Chowdhary, B. Cairns, M. I. Mishchenko, P. V. Hobbs, G. F. Cota, J. Redemann, K. Rutledge, B. N. Holben, and E. Russell, "Retrieval of aerosol scattering and absorption properties from photopolarimetric observations over the ocean during the CLAMS experiment," J. Atmos. Sci. **62**, 1093–1117 (2005).
38. M. Gao, P.-W. Zhai, B. A. Franz, Y. Hu, K. Knobelspiesse, P. J. Werdell, A. Ibrahim, B. Cairns, and A. Chase, "Inversion of multiangular polarimetric measurements over open and coastal ocean waters: a joint retrieval algorithm for aerosol and water-leaving radiance properties," Atmos. Meas. Tech. **12**, 3921–3941 (2019).
39. O. P. Hasekamp, P. Litvinov, and A. Butz, "Aerosol properties over the ocean from PARASOL multiangle photopolarimetric measurements," J. Geophys. Res. Atmos. **116** (2011).
40. J. Chowdhary, B. Cairns, F. Waquet, K. Knobelspiesse, M. Ottaviani, J. Redemann, and L. Travis, "Sensitivity of multiangle, multispectral polarimetric remote sensing over open oceans to water-leaving radiance: analyses of RSP data acquired during the MILAGRO campaign," Remote Sens. Environ. **118**, 284–308 (2012).
41. Y. Fan, W. Li, N. Chen, J.-H. Ahn, Y.-J. Park, S. Kratzer, T. Schroeder, J. Ishizaka, R. Chang, and K. Stamnes, "OC-SMART: a machine learning based data analysis platform for satellite ocean color sensors," Remote Sens. Environ. **253**, 112236 (2021).
42. M. Gao, P.-W. Zhai, B. A. Franz, K. Knobelspiesse, A. Ibrahim, B. Cairns, S. E. Craig, G. Fu, O. Hasekamp, Y. Hu, and P. J. Werdell, "Inversion of multiangular polarimetric measurements from the ACEPOL campaign: an application of improving aerosol property and hyperspectral ocean color retrievals," Atmos. Meas. Tech. **13**, 3939–3956 (2020).
43. M. Gao, B. A. Franz, K. Knobelspiesse, P.-W. Zhai, V. Martins, S. Burton, B. Cairns, R. Ferrare, J. Gales, O. Hasekamp, Y. Hu, A. Ibrahim, B. McBride, A. Puthukkudy, P. J. Werdell, and X. Xu, "Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model," Atmos. Meas. Tech. **14**, 4083–4110 (2021).
44. P.-W. Zhai, Y. Hu, C. R. Trepte, and P. L. Lucker, "A vector radiative transfer model for Coupled Atmosphere and Ocean Systems based on successive order of scattering method," Opt. Express **17**, 2057–2079 (2009).
45. P.-W. Zhai, Y. Hu, J. Chowdhary, C. R. Trepte, P. L. Lucker, and D. B. Josset, "A vector radiative transfer model for Coupled Atmosphere and Ocean Systems with a rough interface," J. Quant. Spectrosc. Radiat. Transfer **111**, 1025–1040 (2010).
46. M. H. Pinkerton, G. F. Moore, S. J. Lavender, M. P. Gall, K. Oubelkheir, K. M. Richardson, P. W. Boyd, and J. Aiken, "A method for estimating inherent optical properties of New Zealand continental shelf waters from satellite ocean colour measurements," New Zealand J. Mar. Freshwater Res. **40**, 227–247 (2006).
47. S. Sathyendranath, L. Proeur, and A. Morel, "A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters," Int. J. Remote Sens. **10**, 1373–1394 (1989).
48. P. Wang, E. S. Boss, and C. Roesler, "Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color," Appl. Opt. **44**, 4074–4085 (2005).
49. R. J. Brewin, S. Sathyendranath, D. Müller, *et al.*, "The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms," Remote Sens. Environ. **162**, 271 (2015).
50. P. J. Werdell, C. S. Roesler, and J. I. Goes, "Discrimination of phytoplankton functional groups using an ocean reflectance inversion model," Appl. Opt. **53**, 4833–4849 (2014).
51. P. J. Werdell and S. W. Bailey, "An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation," Remote Sen. Environ. **98**, 122–140 (2005).
52. A. Valente, S. Sathyendranath, V. Brotas, *et al.*, "A compilation of global bio-optical in situ data for ocean-colour satellite applications—version two," Earth Syst. Sci. Data **11**, 1037–1068 (2019).
53. S. E. Craig, Z. Lee, and K. Du, "Top of atmosphere, hyperspectral synthetic dataset for PACE (phytoplankton, aerosol, and ocean ecosystem) ocean color algorithm development," National Aeronautics and Space Administration, (2020), https://doi.org/10.1594/PANGAEA.915747.
54. I. Cetinic, C. R. McClain, P. J. Werdell, Z. Ahmad, B. A. Franz, E. M. Karakoylu, L. I. McKinna, and F. S. Patt, "PACE technical report series, volume 6: data product requirements and error budgets consensus document," Technical report NASA/TM–2018-219027 (National Aeronautics and Space Administration, 2019).
55. J. Wei, Z. Lee, and S. Shang, "A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments," J. Geophys. Res. Oceans **121**, 8189–8207 (2016).
56. G. Zibordi, F. Mélin, J.-F. Berthon, B. Holben, I. Slutsker, D. Giles, D. D'Alimonte, D. Vandemark, H. Feng, G. Schuster, B. E. Fabbri, S. Kaitala, and J. Seppälä, "AERONET-OC: a network for the validation of ocean color primary products," J. Atmos. Ocean. Technol. **26**, 1634–1651 (2009).
57. H. R. Gordon, O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, "A semianalytic radiance model of ocean color," J. Geophys. Res. Atmos. **93**, 10909–10924 (1988).
58. Z. Lee, S. Shang, Y. Wang, J. Wei, and J. Ishizaka, "Nature of optical products inverted semianalytically from remote sensing reflectance of stratified waters," Limnol. Oceanogr. **65**, 387–400 (2020).
59. Z. Lee, K. L. Carder, R. F. Chen, and T. G. Peacock, "Properties of the water column and bottom derived from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data," J. Geophys. Res. Oceans **106**, 11639–11651 (2001).
60. R. M. Pope and E. S. Fry, "Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements," Appl. Opt. **36**, 8710–8723 (1997).
61. X. Zhang, L. Hu, and M.-X. He, "Scattering by pure seawater: effect of salinity," Opt. Express **17**, 5698–5710 (2009).
62. A. Bricaud, A. Morel, M. Babin, K. Allali, and H. Claustre, "Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: analysis and implications for bio-optical models," J. Geophys. Res. Oceans **103**, 31033 (1998).
63. K. J. Voss, "A spectral model of the beam attenuation coefficient in the ocean and coastal areas," Limnol. Oceanogr. **37**, 501–509 (1992).
64. A. L. Whitmire, E. Boss, T. J. Cowles, and W. S. Pegau, "Spectral variability of the particulate backscattering ratio," Opt. Express **15**, 7019–7031 (2007).
65. A. Morel and B. Gentili, "A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data," Remote Sens. Environ. **113**, 998–1011 (2009).
66. W. H. Slade, E. Boss, and C. Russo, "Effects of particle aggregation and disaggregation on their inherent optical properties," Opt. Express **19**, 7945–7959 (2011).

67. Y. Huot, A. Morel, M. S. Twardowski, D. Stramski, and R. A. Reynolds, "Particle optical backscattering along a chlorophyll gradient in the upper layer of the eastern South Pacific Ocean," Biogeosciences **5**, 495–507 (2008).
68. J. Moré, B. Garbow, and K. Hillström, "User guide for MINPACK-1. [In FORTRAN]," Technical report ANL-80-74 (Argonne National Laboratory, 1980).
69. F. James, "MINUIT: Function minimization and error analysis," Reference manual, version 94.1, (CERN Program Library Long Writeup D506, 1998).
70. W. A. Snyder, R. A. Arnone, C. O. Davis, W. Goode, R. W. Gould, S. Ladner, G. Lamela, W. J. Rhea, R. Stavn, M. Sydor, and A. Weidemann, "Inversion of multiangular polarimetric measurements from the ACEPOL campaign: an application of improving aerosol property and hyperspectral ocean color retrievals," Appl. Opt. **47**, 666–677 (2008).
71. R. R. Sokal and F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman and Company, 1969).
72. J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," Lancet **327**, 307–310 (1986).
73. B. B. Cael, K. Bisson, E. Boss, and Z. K. Ericksons, "How many independent quantities can be extracted from ocean color?" Limnol. Oceangr. (to be published, 2023).
74. F. Xu, D. J. Diner, O. Dubovik, and Y. Schechner, "A correlated multi-pixel inversion approach for aerosol remote sensing," Remote Sens. **11**, 746 (2019).
75. B. Cairns, E. E. Russell, and L. D. Travis, "Research Scanning Polarimeter: calibration and ground-based measurements," Proc. SPIE **3754**, 186–196 (1999).
76. B. Cairns, E. E. Russell, J. D. LaVeigne, and P. M. Tennant, "Research Scanning Polarimeter and airborne usage for remote sensing of aerosols," Proc. SPIE **5158**, 33–44 (2003).
77. J. V. Martins, T. Nielsen, C. Fish, L. Sparr, R. Fernandez-Borda, M. Schoeberl, and L. Remer, "HARP CubeSat–an innovative hyperangular imaging polarimeter for earth science applications," *Small Sat Pre-Conference Workshop*, Logan, Utah, USA (2014), Vol. **20**.
78. M. Smit, J. Rietjens, O. P. Hasekamp, A. Di Noia, G. van Harten, B. E. Rheingans, D. J. Diner, F. C. Seidel, and O. V. Kalashnikova, "First results of the SPEX airborne multi-angle spectropolarimeter-aerosol retrievals over ocean and intercomparison with AirMSPI," in *AGU Fall Meeting Abstracts* (2016), paper A44A-06.
79. J. M. Smit, J. H. Rietjens, A. di Noia, O. P. Hasekamp, W. Laauwen, B. Cairns, B. van Diedenhoven, and A. Wasilewski, "In-flight validation of SPEX airborne spectro-polarimeter onboard NASA's research aircraft ER-2," Proc. SPIE **11180**, 111800N (2019).

130.85.194.150 (ge 2023-05-18 13:14:06)