

# Generalized ocean color inversion model for retrieving marine inherent optical properties

P. Jeremy Werdell,<sup>1,2,\*</sup> Bryan A. Franz,<sup>1</sup> Sean W. Bailey,<sup>1,3</sup> Gene C. Feldman,<sup>1</sup>  
Emmanuel Boss,<sup>2</sup> Vittorio E. Brando,<sup>4</sup> Mark Dowell,<sup>5</sup> Takafumi Hirata,<sup>6</sup>  
Samantha J. Lavender,<sup>7</sup> ZhongPing Lee,<sup>8</sup> Hubert Loisel,<sup>9</sup>  
Stéphane Maritorena,<sup>10</sup> Frédéric Mélin,<sup>5</sup> Timothy S. Moore,<sup>11</sup>  
Timothy J. Smyth,<sup>12</sup> David Antoine,<sup>13</sup> Emmanuel Devred,<sup>14</sup>  
Odile Hembise Fanton d'Andon,<sup>15</sup> and Antoine Mangin<sup>15</sup>

<sup>1</sup>NASA Goddard Space Flight Center, Code 616, Greenbelt, Maryland 20771, USA

<sup>2</sup>School of Marine Sciences, University of Maine, Orono, Maine 04401, USA

<sup>3</sup>Futuretech Corporation, Greenbelt, Maryland 20770, USA

<sup>4</sup>CSIRO Land & Water, Environmental Earth Observation Program, G.P.O. Box 1666, Canberra, Australia

<sup>5</sup>Joint Research Centre, European Commission, Ispra 21027, Italy

<sup>6</sup>Faculty of Environmental Earth Science, Hokkaido University, Sapporo 060-0810, Japan

<sup>7</sup>Pixalytics Ltd., Plymouth PL6 8BX, UK

<sup>8</sup>Environmental Earth and Ocean Sciences, University of Massachusetts at Boston, Boston, Massachusetts 02125, USA

<sup>9</sup>Laboratoire d'Océanologie et de Géosciences, Université du Littoral—Côte d'Opale, Wimereux F-62930, France

<sup>10</sup>Earth Research Institute, University of California Santa Barbara, Santa Barbara, California 93106, USA

<sup>11</sup>Ocean Process Analysis Laboratory, University of New Hampshire, Durham, New Hampshire 03824, USA

<sup>12</sup>Plymouth Marine Laboratory, Plymouth PL1 3DH, UK

<sup>13</sup>CNRS, Laboratoire d'Océanographie de Villefranche, BP 08 06238 Villefranche sur mer, France

<sup>14</sup>Unité mixte internationale Takuvik/Joint international Laboratory, Université Laval, Québec GV1 0A6, Canada

<sup>15</sup>ACRI-ST, 06904 Sophia Antipolis Cedex, France

\*Corresponding author: jeremy.werdell@nasa.gov

Received 20 December 2012; revised 7 February 2013; accepted 11 February 2013;  
posted 14 February 2013 (Doc. ID 182094); published 22 March 2013

Ocean color measured from satellites provides daily, global estimates of marine inherent optical properties (IOPs). Semi-analytical algorithms (SAAs) provide one mechanism for inverting the color of the water observed by the satellite into IOPs. While numerous SAAs exist, most are similarly constructed and few are appropriately parameterized for all water masses for all seasons. To initiate community-wide discussion of these limitations, NASA organized two workshops that deconstructed SAAs to identify similarities and uniqueness and to progress toward consensus on a unified SAA. This effort resulted in the development of the generalized IOP (GIOP) model software that allows for the construction of different SAAs at runtime by selection from an assortment of model parameterizations. As such, GIOP permits isolation and evaluation of specific modeling assumptions, construction of SAAs, development of regionally tuned SAAs, and execution of ensemble inversion modeling. Working groups associated with the

workshops proposed a preliminary default configuration for GIOP (GIOP-DC), with alternative model parameterizations and features defined for subsequent evaluation. In this paper, we: (1) describe the theoretical basis of GIOP; (2) present GIOP-DC and verify its comparable performance to other popular SAAs using both *in situ* and synthetic data sets; and, (3) quantify the sensitivities of their output to their parameterization. We use the latter to develop a hierarchical sensitivity of SAAs to various model parameterizations, to identify components of SAAs that merit focus in future research, and to provide material for discussion on algorithm uncertainties and future ensemble applications. © 2013 Optical Society of America

OCIS codes: 010.4450, 280.4991.

## 1. Introduction

Satellite ocean color instruments, such as the NASA Moderate Resolution Imaging Spectroradiometer on-board Aqua (MODISA) and ESA Medium Resolution Imaging Spectrometer (MERIS), provide daily global estimates of marine inherent optical properties (IOPs). Remotely sensed IOPs, namely the spectral absorption and backscattering coefficients of the surface water column and its particulate and dissolved constituents, describe the contents of the upper ocean, information critical to furthering scientific understanding of biogeochemical processes, such as carbon exchanges, phytoplankton biodiversity shifts, and responses to climatic disturbances. As such, the international community has invested significant effort in ensuring and improving the quality of remotely sensed IOPs. Recognizing this, the International Ocean Colour Coordinating Group (IOCCG) established a working group to compile and compare popular methods for estimating IOPs from ocean color [1]. NASA recently extended this effort through a series of internationally attended workshops aimed at: (a) deconstructing each method to identify similarities and uniqueness; (b) identifying strategies to provide uncertainties for each method; and (c) achieving community-wide consensus on a unified method for generating global satellite IOP products [2].

Semi-analytical algorithms (SAAs) provide one mechanism for inverting ocean color into IOPs through a combination of empiricism and radiative transfer theory. Many published SAAs attempt to retrieve spectral IOPs from spectral remote-sensing reflectances [ $R_{rs}(\lambda)$ ;  $\text{sr}^{-1}$ ] and to further separate the total absorption and backscattering coefficients into subcomponents [e.g., the absorption of algal, nonalgal, and colored dissolved organic matter (CDOM)] [3–11]. These SAAs generally assume spectral shape functions (or eigenvectors) of the constituent absorption and scattering components and retrieve the magnitudes (or eigenvalues) of each constituent required to match the spectral distribution of  $R_{rs}(\lambda)$ . Thus, SAAs often differ only in the assumptions employed to define the eigenvectors and in the mathematical methods applied to calculate the eigenvalues. In lieu of this, and in support of the aforementioned workshops, the NASA Ocean Biology Processing Group (OBPG) [12] developed

the generalized IOP (GIOP) framework to facilitate controlled evaluation of the varied SAA approaches within a satellite data processing environment [13].

Briefly, GIOP allows construction of different IOP models at runtime by selection from a wide assortment of published absorption and backscattering eigenvectors [4,6–8,14–17]. As such, GIOP permits isolation and evaluation of specific modeling assumptions, construction of new SAAs, development of regionally tuned SAAs (e.g., [18]), and ensemble inversion modeling. The OBPG implemented GIOP is standard NASA ocean color processing code [19] and currently distributes GIOP to the research community via the Sea-viewing Wide Field-of-View Sensor (SeaWiFS) Data Analysis System (SeaDAS) [20]. The working group associated with the NASA GIOP workshops (October 2008 and September 2010) [2] proposed the preliminary configuration of GIOP, with alternative model parameterizations and features defined for subsequent evaluation. In this paper, we: (a) describe the theoretical basis of GIOP; (b) present its preliminary configuration; (c) evaluate and validate this configuration; and (d) quantify the sensitivities of output eigenvalues to their assumed eigenvectors. We provide a brief summary of the model components within GIOP as compared to several common SAAs in Appendix A.

## 2. Methods

### A. Model Development

Ocean color satellite instruments provide estimates of  $R_{rs}(\lambda)$  after atmospheric correction. The method of Lee *et al.* [7] provides a convenient method to convert  $R_{rs}(\lambda)$  to their subsurface values:

$$r_{rs}(\lambda, 0^-) = \frac{R_{rs}(\lambda)}{0.52 + 1.7R_{rs}(\lambda)}. \quad (1)$$

Subsurface remote-sensing reflectances relate to marine IOPs via

$$r_{rs}(\lambda, 0^-) = G_1(\lambda)u(\lambda) + G_2(\lambda)u(\lambda)^2, \quad (2)$$

where the two  $G(\lambda)$  ( $\text{sr}^{-1}$ ) vary with illumination conditions and geometry, sea surface properties,

bidirectional effects, and the shape of the marine volume scattering function. Most often,  $u(\lambda)$  is described as

$$u(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}, \quad (3)$$

where  $b_b$  is the total backscattering coefficient ( $\text{m}^{-1}$ ) and  $a$  is the total absorption coefficient ( $\text{m}^{-1}$ ) [21], but formulations such as  $u(\lambda) = b_b(\lambda)/a(\lambda)$  have also been adopted [22]. Common methods for estimating  $G(\lambda)$  include Gordon *et al.* [21], where  $G_1$  and  $G_2$  are spectrally fixed to 0.0949 and 0.0794 (see [7,23] for alternative coefficients), and the tabulated results of Morel *et al.* [22], where  $G_1$  is estimated using solar and sensor geometries and an estimate of algal biomass and  $G_2$  is set to 0. GIOP supports all of these options. In practice, most SAAs first solve for  $u(\lambda)$ , then decompose (or, invert)  $u(\lambda)$  into its component IOPs.

The absorption coefficient can be expanded as the sum of all absorbing components. Further, each component can be expressed as the product of its concentration-specific absorption spectrum (eigenvector;  $a^*$ ) and its concentration or amplitude (eigenvalue;  $A$ ):

$$a(\lambda) = a_w(\lambda) + \sum_{i=1}^{N_\phi} A_{\phi_i} a_{\phi_i}^*(\lambda) + \sum_{i=1}^{N_d} A_{d_i} a_{d_i}^*(\lambda) + \sum_{i=1}^{N_g} A_{g_i} a_{g_i}^*(\lambda), \quad (4)$$

where the subscripts  $w$ ,  $\phi$ ,  $d$ , and  $g$  indicate contributions by water, phytoplankton, nonalgal particles (NAP), and CDOM. Both  $a_d^*(\lambda)$  and  $a_g^*(\lambda)$  are commonly expressed as

$$a_{d,g}^*(\lambda) = \exp(-S_{d,g}\lambda), \quad (5)$$

where  $S_d$  and  $S_g$  typically vary between 0.01 and 0.02  $\text{nm}^{-1}$  in natural waters [24]. As the spectral shapes of NAP and CDOM absorption differ only in their exponential slopes, the two components are typically combined for satellite applications and Eq. (4) becomes

$$a(\lambda) = a_w(\lambda) + \sum_{i=1}^{N_\phi} A_{\phi_i} a_{\phi_i}^*(\lambda) + \sum_{i=1}^{N_{dg}} A_{dg_i} a_{dg_i}^*(\lambda), \quad (6)$$

which is the default option for GIOP.

Total backscattering can be expanded to

$$b_b(\lambda) = b_{bw}(\lambda) + \sum_{i=1}^{N_{bp}} B_{bp_i} b_{bp_i}^*(\lambda), \quad (7)$$

where the subscripts  $bw$  and  $bp$  indicate contributions by seawater and particles.  $B_{bp}$  provides the

eigenvalue and a power function often represents the eigenvector:

$$b_{bp}^*(\lambda) = \lambda^{S_{bp}}, \quad (8)$$

where  $S_{bp}$  typically varies between  $-2$  and  $0$  from small to large particles. While commonly employed in the remote-sensing paradigm, we acknowledge the validity of the power function for  $b_{bp}^*(\lambda)$  remains debatable [25–27]. Both  $a_w(\lambda)$  and  $b_{bw}(\lambda)$  are known [28,29].

Using  $R_{rs}(\lambda)$  and eigenvectors as input, eigenvalues for absorption ( $A$ ) and backscattering ( $B$ ) can be estimated via linear or nonlinear least squares inversion of Eqs. (1)–(3). For example, Roesler and Perry [3] used the nonlinear optimization scheme of Levenberg–Marquardt (LM), while Hoge and Lyon [4] showed the problem could be linearized and thus directly solved via linear matrix inversion. GIOP supports both linear and nonlinear optimization and matrix inversion schemes [13]. The optimized eigenvalues represent the relative contributions of each defined absorbing and scattering constituent. In the special case where  $a_\phi^*(\lambda)$  is provided as chlorophyll-specific absorption ( $\text{m}^2 \text{mg}^{-1}$ ), the eigenvalue  $A_\phi$  provides an estimate of the chlorophyll concentration,  $C_a$  ( $\text{mgm}^{-3}$ ). Note that this model describes each component of absorption and backscattering as a linear sum of subcomponents, presumably with unique spectral dependencies [symbolized by the summation over  $N$  in Eqs. (4), (6), and (7)]. In this way, the absorption characteristics for different phytoplankton populations and the scattering characteristics of multiple size distributions of suspended particles can be represented, or Eq. (6) can be re-expanded to Eq. (4).

## B. Model Configuration

GIOP provides a satellite data processing framework within which an SAA can be constructed at runtime by selection from an assortment of published eigenvectors (Table 1). Franz and Werdell [13] provide a detailed description of its use within the SeaDAS environment [20]. The general form of the GIOP inversion model (Section 2.A) is common to a number of published approaches (e.g., [3–5,8,9,30]) whose differences reside in the choice of eigenvectors employed, the number of eigenvalues resolved, the optimization method selected, and the number of sensor wavelengths considered in the optimization. A unique instance of GIOP is therefore defined by specifying eigenvectors for each optically significant constituent assumed to exist in the water column. The working groups associated with the NASA GIOP workshops [2] proposed the following preliminary, default configuration of GIOP, hereafter referred to as GIOP default configuration (GIOP-DC), to support NASA's production and distribution of global IOP data products:

Table 1. Summary of Eigenvectors Available for Use in GIOP (as of March 2011)\*

Eigenvector	Description	Reference
$a_\phi^*(\lambda)$	User-provided $a_\phi^*(\lambda)$	
	Maritorena <i>et al.</i> (2002) tabulated $a_\phi^*(\lambda)$	[8]
	<b>Bricaud <i>et al.</i> (1998)-derived <math>a_\phi^*(\lambda)</math> using OC-derived <math>C_a</math></b>	[14]
	Ciotti and Bricaud (2006)-derived $a_\phi^*(\lambda)$ using user-provided size fraction	[17]
$a_{dg}^*(\lambda)$	<b>Eq. (5) with user-provided <math>S_{dg}</math></b>	
	Eq. (5) with Lee <i>et al.</i> (2002)-derived $S_{dg}$	[7]
	Eq. (5) with Franz and Werdell (2010)-derived $S_{dg}$	[13]
	User-provided $a_{dg}^*(\lambda)$	
	Eq. (8) with user-provided $S_{bp}$	
$b_{bp}^*(\lambda)$	Eq. (8) with Hoge and Lyon (1996)-derived $S_{bp}$	[4]
	<b>Eq. (8) with Lee <i>et al.</i> (2002)-derived <math>S_{bp}</math></b>	[7]
	Eq. (8) with Ciotti <i>et al.</i> (1999)-derived $S_{bp}$	[15]
	Eq. (8) with Morel and Maritorena (2001)-derived $S_{bp}$	[22]
	Eq. (8) with Loisel and Stramski (2000)-derived $S_{bp}$	[6]
	User-provided $b_{bp}^*(\lambda)$	
	Loisel and Stramski (2000)-derived $b_{bp}^*(\lambda)$	[6]
Lee <i>et al.</i> (2002)-derived $b_{bp}^*(\lambda)$	[7]	

\*Boldface indicates the eigenvector used in GIOP-DC.

- Relate  $r_{rs}(\lambda, 0^-)$  to IOPs using Eq. (2) and

$$u(\lambda) = \frac{b_{bw}(\lambda) + B_{bp}b_{bp}^*(\lambda)}{b_{bw}(\lambda) + B_{bp}b_{bp}^*(\lambda) + a_w(\lambda) + A_{dg}a_{dg}^*(\lambda) + A_\phi a_\phi^*(\lambda)}. \quad (9)$$

- Spectrally invariant  $G_1$  and  $G_2$  from Gordon *et al.* [21]
- $b_{bp}^*(\lambda)$  from Eq. (8) and  $S_{bp}$  from Lee *et al.* [7]
- $a_{dg}^*(\lambda)$  from Eq. (5) and  $S_{dg} = 0.018 \text{ nm}^{-1}$  [24]
- $a_\phi^*(\lambda)$  from Bricaud *et al.* [14], estimated using OC-derived  $C_a$  [31] and normalized such that  $a_\phi^*(443) = 0.055 \text{ m}^2 \text{ mg}^{-1}$ .

The latter normalization allows the spectral shape of  $a_\phi^*(\lambda)$  to change with an estimate of  $C_a$ , but constrains its magnitude to an average value for oceanic water [3,8,14].

Our choice of default parameterizations served two purposes: to provide some consistency with previously developed SAAs and to acknowledge the emerging quality of variable, dynamically selected eigenvectors. In particular, we adopted  $G_1$  and  $G_2$  from Gordon *et al.* [21] to be consistent with previously published work [3–5,7–9] and variable  $S_{bp}$  and  $a_\phi^*(\lambda)$  to better represent heterogeneous natural environments than would be possible with fixed values. Our choice of a fixed  $S_{dg}$  represents a compromise between multiple published values [8,24], although we expect GIOP-DC to ultimately adopt a variable  $S_{dg}$  as methods for its dynamic calculation improve. Note, we did not make these choices based on comparisons with *in situ* or synthetic data sets, nor do we suggest that these parameterizations represent all natural oceanic conditions at all times. Our choices simply represent a consistent, modern SAA configuration from which to evolve as our understanding of marine bio-optics improves.

In GIOP-DC, the three unknown eigenvalues ( $B_{bp}$ ,  $A_{dg}$ , and  $A_\phi$ ) are optimized using LM, and wavelengths between 400 and 700 nm are considered. Our convergence criterion tests both absolute and relative changes in the optimized eigenvalues between iterations of the LM algorithm. Convergence is reached when the change in any eigenvalue,  $X$ , between iteration  $i$  and iteration  $i + 1$  satisfies  $|X_{i+1} - X_i| < 0.0001 + 0.0001|X_i|$ . If the number of iterations exceeds 50, the iteration is stopped and the result is treated as algorithm failure. Note that these criteria do not require the results to be valid or even that the model provides a good fit to the data. Despite this, we consider the IOP retrievals to be valid when

- $-0.05b_{bw}(\lambda) \leq b_{bp}(\lambda) \leq 0.05 \text{ m}^{-1}$
- $-0.05a_w(\lambda) \leq a_{dg}(\lambda) \leq 5 \text{ m}^{-1}$
- $-0.05a_\phi(\lambda) \leq a_\phi(\lambda) \leq 5 \text{ m}^{-1}$
- $\Delta R_{rs} \leq 33\%$ ,

where the mean relative  $R_{rs}(\lambda)$  difference, a measure of goodness of fit, is derived as

$$\Delta R_{rs} = \frac{100\%}{N_\lambda} \sum_{i=1}^{N_\lambda} \frac{|\hat{R}_{rs}(\lambda_i) - R_{rs}(\lambda_i)|}{R_{rs}(\lambda_i)}, \quad (10)$$

for  $400 \leq \lambda \leq 600 \text{ nm}$  with  $\hat{R}_{rs}$  indicating remote-sensing reflectance reconstructed from the modeled IOPs. We consider slightly negative IOP values to be viable (rather, statistically equivalent to zero), as the retrieved eigenvalues include some uncertainty. We provide a brief comparison of GIOP-DC (and GIOP, in general) and several other SAAs commonly used in satellite ocean color applications in Appendix A.

### C. Uncertainties

The LM optimization method minimizes a  $\chi^2$  merit function defined as

$$\chi^2 = \sum_{i=1}^{N_\lambda} \frac{(\hat{R}_{rs}(\lambda_i) - R_{rs}(\lambda_i))^2}{\sigma^2(\lambda_i)}, \quad (11)$$

where  $\sigma(\lambda_i)$  are the input uncertainties on  $R_{rs}(\lambda_i)$ . LM makes use of the Jacobian matrix ( $J$ ), which is derived by evaluating the partial derivatives of the merit function with respect to the optimized eigenvalues at each  $\lambda_i$ . From the Jacobian, the covariance matrix ( $M$ ) is then constructed as  $M = (J^T J)^{-1}$ , resulting in a square matrix with dimensionality equal to the number of model eigenvalues. The square-root of the diagonal elements of  $M$  represent the uncertainty on the optimized eigenvalues, taking into account both the weighting and scale of the input uncertainties. If reliable values of  $\sigma(\lambda_i)$  are not available, they are set to 1.0 and the optimization is unweighted. In that case, the uncertainty is computed using the square-root of the diagonal terms of the variance–covariance matrix,  $\sigma^2 M$ , where

$$\sigma^2 = \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} (\hat{R}_{rs}(\lambda_i) - R_{rs}(\lambda_i))^2. \quad (12)$$

The GIOP-DC configuration of GIOP currently utilizes an unweighted optimization with the variance–covariance matrix utilized to estimate uncertainties on the model eigenvalues. The uncertainty on the derived IOPs is then determined by multiplying the uncertainty of the optimized eigenvalue by the associated eigenvector.

#### D. Data Acquisition

We acquired coincident observations of *in situ* and synthesized  $R_{rs}(\lambda)$ ,  $b_{bp}(\lambda)$ ,  $a_{dg}(\lambda)$ ,  $a_\phi(\lambda)$ , and  $C_a$  from the NASA bio-Optical Marine Algorithm Data set (NOMAD) [32] and the IOCCG Ocean Colour Algorithms Working Group synthetic data set [1]. We also acquired coincident Level-2 satellite-to-*in situ* match-ups for the NASA SeaWiFS and MODISA instruments from the OBP (http://seabass.gsfc.nasa.gov/seabasscgi/validation\_search.cgi). Satellite data processing and quality assurance for the match-ups followed Bailey and Werdell [33]. Specifically: (1) temporal coincidence was defined as  $\pm 3$  h; (2) satellite values were the filtered median (via the semi-interquartile range) of all unmasked pixels in a  $5 \times 5$  box centered on the *in situ* target; and (3) satellite values were excluded when the coefficient of variation for the given product within this box exceeded 0.15. SeaWiFS and MODISA data were processed following their R2010.0 (September 2010) and R2012.0 (May 2012) reprocessing configurations, respectively [34]. The *in situ* data used in the match-up analyses include a small subset of NOMAD, plus additional  $R_{rs}(\lambda)$ ,  $C_a$ ,  $b_{bp}(\lambda)$ ,  $a_{dg}(\lambda)$ , and  $a_\phi(\lambda)$  archived in the NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS) [35]. Sample sizes for  $b_{bp}(\lambda)$  differ from  $a(\lambda)$ ,  $a_{dg}(\lambda)$ , and  $a_\phi(\lambda)$  in NOMAD and the match-up data sets because of instrumental variability in field campaigns included in these compilations.

#### E. Validation Analyses

We adopted a validation approach based on model-measurement regression statistics and a spectral goodness of fit metric. We compared modeled and ground-truth IOPs, with the former calculated using *in situ*, synthesized, and satellite  $R_{rs}(\lambda)$  from the NOMAD, IOCCG, and satellite-to-*in situ* match-up data sets. We generated Type II linear regression statistics and estimates of  $\Delta$ IOP for all of the above, with  $\Delta$ IOP defined as

$$\Delta\text{IOP} = \frac{200\%}{N_\lambda} \sum_{i=1}^{N_\lambda} \frac{|\hat{\text{IOP}}(\lambda_i) - \text{IOP}(\lambda_i)|}{\hat{\text{IOP}}(\lambda_i) + \text{IOP}(\lambda_i)}, \quad (13)$$

for  $400 \leq \lambda \leq 600$  nm. IOP and  $\hat{\text{IOP}}$  indicate a measured and modeled absorption or backscattering coefficient, respectively. Accordingly, we report  $\Delta$ IOP elsewhere as  $\Delta b_{bp}$ ,  $\Delta a$ ,  $\Delta a_{dg}$ , and  $\Delta a_\phi$ . We generated modeled IOPs using *in situ*  $R_{rs}(\lambda)$  and calculated  $\Delta R_{rs}$  and  $\Delta$ IOP distribution medians and semi-interquartile ranges for various populations (here, the semi-interquartile range is the range covered by values such that 50% occur with equal probability on either side of the median). Furthermore, for the analysis of these delta statistics, we stratified the ground-truth stations into three trophic levels: oligotrophic water has  $C_a \leq 0.1$  mg m<sup>-3</sup>; mesotrophic water has  $0.1 < C_a \leq 1$  mg m<sup>-3</sup>; and eutrophic water has  $C_a > 1$  mg m<sup>-3</sup>. Table 2 provides calculations for other comparative statistics.

#### F. Sensitivity Analyses

We executed 12 additional unique instances of GIOP to evaluate the sensitivity of the GIOP-DC configuration to alternate selections of eigenvectors. Each of the 12 instances included a single alternate parameterization: (1,2)  $S_{dg} \pm 33\%$  ( $= 0.012$  and  $0.024$  nm<sup>-1</sup>), (3)  $S_{dg}$  dynamically calculated using Lee *et al.* [7], (4,5)  $S_{bp}$  from Lee *et al.* [7]  $\pm 33\%$ ; (6,7) OC-derived  $C_a \pm 33\%$  prior to input into Bricaud *et al.* [14]; (8)  $a_\phi^*(\lambda)$  from Bricaud *et al.* [14] with  $C_a$  fixed at  $0.18$  mg m<sup>-3</sup>; (9)  $a_\phi^*(\lambda)$  from Ciotti and Bricaud [17] with a size fraction of 0.5; (10)  $G(\lambda)$  from Morel *et al.* [22], (11) optimization using linear matrix inversion, and (12) optimization considering only  $400 \leq \lambda \leq 600$  nm. We quantified spectral changes in modeled IOPs for each alternate parameterization (relative to the baseline GIOP-DC configuration) using Type II regression statistics, estimates of  $\Delta$ IOP, and Taylor [36] and Target [37] summary diagrams. The latter provide convenient means for simultaneously considering magnitudes of deviations, correlations, and biases between the alternate runs and GIOP-DC. To minimize the impact of uncertainties inherent to the *in situ* IOP measurements, we only considered the synthetic IOCCG data set in these analyses (note, this does not remove uncertainties in the truth data set, but rather, limits them to those associated with the inherent assumptions used to generate this synthetic data set).

Table 2. Regression Statistics for GIOP-DC Using the NOMAD and IOCCG Data Sets<sup>a</sup>

	NOMAD						IOCCG				
	<i>N</i>	<i>r</i> <sup>2</sup>	Slope (SE)	Ratio	MPD	<i>N</i>	<i>r</i> <sup>2</sup>	Slope (SE)	Ratio	MPD	
<i>b<sub>bp</sub></i>	412	217	0.49	0.99 (0.05)	1.26	28.2	437	0.92	1.08 (0.01)	0.98	10.1
	443	217	0.52	1.00 (0.05)	1.27	28.8	437	0.93	1.07 (0.01)	1.00	9.3
	490	217	0.57	1.04 (0.05)	1.29	30.0	437	0.94	1.07 (0.01)	0.99	8.2
	555	217	0.60	1.06 (0.05)	1.30	32.8	437	0.95	1.09 (0.01)	0.91	11.8
	670	217	0.64	1.09 (0.05)	1.27	31.3	437	0.96	1.09 (0.01)	0.90	13.0
<i>α</i>	412	649	0.90	1.14 (0.01)	0.89	20.5	421	0.99	1.03 (0.00)	1.03	7.2
	443	657	0.90	1.13 (0.01)	0.88	21.8	427	0.99	1.03 (0.01)	0.99	6.1
	490	657	0.88	1.13 (0.02)	0.84	24.6	432	0.98	1.03 (0.01)	0.98	8.7
	555	654	0.81	1.21 (0.02)	0.77	37.4	436	0.95	1.11 (0.01)	0.73	29.9
	670	609	0.80	1.18 (0.02)	1.11	37.4	424	0.86	1.01 (0.02)	1.50	58.6
<i>α<sub>dg</sub></i>	412	654	0.83	1.12 (0.02)	0.84	28.8	426	0.98	1.02 (0.01)	1.01	13.1
	443	662	0.81	1.10 (0.02)	0.78	35.1	429	0.97	1.04 (0.01)	0.90	15.9
	490	662	0.76	1.06 (0.02)	0.65	43.2	435	0.95	1.07 (0.01)	0.80	25.0
	555	656	0.71	1.03 (0.02)	0.53	52.7	437	0.91	1.08 (0.02)	0.58	44.1
	670	621	0.61	0.96 (0.03)	0.32	69.8	426	0.82	1.05 (0.02)	0.36	64.2
<i>α<sub>φ</sub></i>	412	676	0.76	1.23 (0.02)	0.94	27.5	419	0.81	1.18 (0.03)	1.16	32.8
	443	682	0.76	1.21 (0.02)	0.98	25.9	419	0.79	1.13 (0.03)	1.20	30.3
	490	682	0.76	1.23 (0.02)	1.08	27.2	422	0.78	1.13 (0.03)	1.34	41.4
	555	673	0.78	1.22 (0.02)	1.25	42.5	422	0.84	1.10 (0.02)	1.36	52.5
	670	680	0.82	1.12 (0.02)	1.33	43.3	422	0.81	1.00 (0.02)	2.01	100.8

<sup>a</sup>*N*, *r*<sup>2</sup>, and Slope (SE) are the sample size, coefficient of determination, and regression slope (standard error of the regression slope), respectively. Ratio is the median ratio calculated as median( $\hat{X}_i/X_i$ ) and MPD is the median percent difference calculated as median( $100\% * |\hat{X}_i/X_i - 1|$ ), with  $\hat{X}_i$  and  $X_i$  indicating each modeled and measured IOP, respectively. With the exception of Ratio and MPD, data were log-transformed.

### 3. Results

Direct comparisons of modeled and ground-truth IOPs provided estimates of the accuracy and precision of GIOP-DC. Overall, spectral IOPs from GIOP-DC (indicated by a carot hat below) reproduced the *in situ* and synthetic values moderately well over the full dynamic range of the NOMAD and IOCCG data sets, with *r*<sup>2</sup> ranging from 0.49 to 0.90 and 0.78 to 0.99, respectively (Fig. 1, Table 2). With the exception of  $\hat{\alpha}_\phi(\lambda)$ , median percent differences (MPD; the caption for Table 2 provides its calculation) for NOMAD showed greater overall variability than for the IOCCG data set, with values ranging largely from 20%–40% compared to 10%–30%. Both data sets showed less variability for  $\hat{b}_{bp}(\lambda)$  and  $\hat{\alpha}(\lambda)$  than for the component absorption products,  $\hat{\alpha}_{dg}(\lambda)$  and  $\hat{\alpha}_\phi(\lambda)$  [38,39]. As demonstrated by their median ratios,  $\hat{b}_{bp}(\lambda)$  and  $\hat{\alpha}(\lambda)$  for NOMAD showed high and low biases (over and underestimates), respectively, while both approached unity for the IOCCG data set. Relative to the ratios for  $\hat{\alpha}(\lambda)$ , ratios for  $\hat{\alpha}_{dg}(\lambda)$  and  $\hat{\alpha}_\phi(\lambda)$  from both data sets showed low and high biases, respectively, although the IOCCG results better demonstrated this behavior. Interestingly, the absolute magnitudes of the biases in component absorption products increased with increasing wavelength for both data sets, possibly due to the increasing contribution of  $\alpha_w(\lambda)$  to  $\alpha(\lambda)$  with increasing wavelength. The magnitudes of the regression slopes for  $\hat{b}_{bp}(\lambda)$  and  $\hat{\alpha}_{dg}(\lambda)$  from

NOMAD also showed strong, albeit opposite, spectral dependency.

Not surprisingly, comparisons of modeled and *in situ* IOPs for SeaWiFS and MODISA resulted in similar statistics to those for NOMAD, as the *in situ* IOPs used in the NOMAD and satellite comparisons stemmed from similar sources (Fig. 2, Table 3). This also implies, however, that the satellite and *in situ*  $R_{rs}(\lambda)$  carry similar spectral uncertainties or that the inversion process remains somewhat insensitive to uncertainties in  $R_{rs}(\lambda)$ . The SeaWiFS and MODISA samples sizes and dynamic ranges were far lower than those for the NOMAD and IOCCG comparisons, yet their MPD were higher, at least for the absorption coefficients. Similar to the NOMAD and IOCCG results, MPD for  $\hat{\alpha}_{dg}(\lambda)$  and  $\hat{\alpha}_\phi(\lambda)$  exceeded those for  $\hat{b}_{bp}(\lambda)$  and  $\hat{\alpha}(\lambda)$ . The median ratios for  $\hat{\alpha}_{dg}(\lambda)$  and  $\hat{\alpha}_\phi(\lambda)$  showed definitive spectral dependency, but only partial low and high biases relative to the ratios for  $\hat{\alpha}(\lambda)$ . Similar to the NOMAD and IOCCG results, the magnitudes of the slopes for  $\hat{b}_{bp}(\lambda)$  and  $\hat{\alpha}_{dg}(\lambda)$  for SeaWiFS showed spectral dependencies. Such behavior was not obvious for MODISA, presumably given its small sample size for  $\hat{\alpha}_{dg}(\lambda)$ . As for the NOMAD and IOCCG results, GIOP-DC appeared to best retrieve  $\hat{\alpha}(\lambda)$ , echoing the findings of an IOCCG working group [1].

GIOP-DC ran successfully on 90% of stations in NOMAD and the IOCCG data set, independent of trophic level (Table 4). The 10% failure rate resulted from a combination of  $\Delta R_{rs} > 33\%$  and nonconvergence of

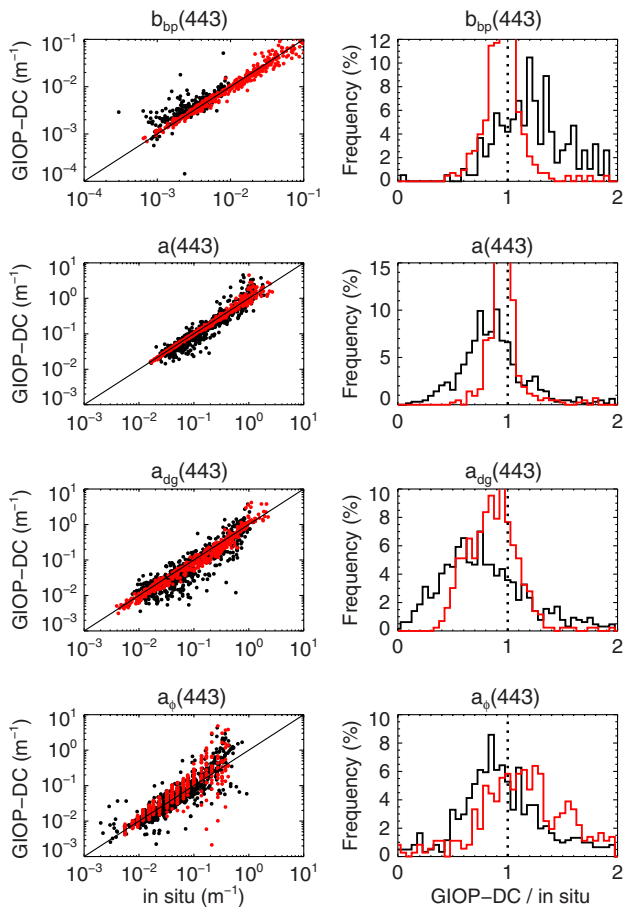


Fig. 1. Comparison of GIOP-DC and ground truth (*in situ*) IOPs at 443 nm from NOMAD (black) and the IOCCG data set (red). The left column shows scatter plots for regression analyses. The right column shows ratios of GIOP-DC to ground truth. See Table 2 for accompanying statistics.

the inversion. When successful,  $\Delta R_{rs}$  fell below 2% on average for all trophic levels for both data sets (Fig. 3). While this indicates the eigenvectors can accurately reproduce  $R_{rs}(\lambda)$ , it does not necessarily suggest that GIOP-DC employs ideal eigenvectors. In contrast to  $\Delta R_{rs}$ ,  $\Delta IOPs$  fell between 25%–50% for NOMAD (excluding the oligotrophic subset) and 5%–35% for the IOCCG data set (excluding the eutrophic,  $\Delta a_\phi$ , Table 4). As will be discussed later, some variability in the NOMAD results stems from the *in situ* data, most of which fail to accurately achieve radiometric closure between  $R_{rs}(\lambda)$  and IOPs because of  $G(\lambda)$  or measurement error [see Eq. (2)]. The IOCCG data set, however, maintains radiometric closure by design and reported  $\Delta b_{bp}$  and  $\Delta a$  distribution means and modes below 15% (Fig. 4). As for the match-up results, delta values for the component absorption products,  $\Delta a_{dg}$  and  $\Delta a_\phi$  were degraded relative to  $\Delta a$ . The latter is true for the full IOCCG data set and for the meso- and eutrophic subsets of NOMAD, which dominate its sampling distribution. For the IOCCG data set,  $\Delta a_{dg}$  remained consistent for the three trophic levels ( $28\% \pm 3\%$ ), while  $\Delta a_\phi$  increased with increasing trophic level. This perhaps

resulted from the decreasing relative concentration of  $a_\phi(\lambda)$  to total absorption. The oligo-, meso-, and eutrophic subsets of the IOCCG data set maintain median ratios for  $a_\phi(443)$  to  $a_\phi(443) + a_{dg}(443)$  of 0.40, 0.31, and 0.23, respectively.

A hierarchy emerges from the sensitivity analyses performed on the IOCCG data set using alternate parameterizations of eigenvectors. Under- and overestimation of  $S_{bp}$  produced eigenvalues that differed from GIOP-DC by only 2%–8% (MPD), with slightly elevated  $\Delta IOPs$  (Table 5). This indicates the retrieved eigenvalues to be fairly insensitive to  $S_{bp}$ , particularly given that GIOP-DC employs a dynamic calculation of this eigenvector [7]. Modifying the  $C_a$  used as input into Bricaud *et al.* [14] to retrieve  $a_\phi^*(\lambda)$  produced eigenvalues that differed from GIOP-DC by only 1%–7% (MPD), with very comparable  $\Delta IOPs$ . We suspect this results from stability in  $a_\phi^*(\lambda)$  from the Bricaud model over narrow ranges of  $C_a$  ( $\pm 33\%$ ). Not surprisingly, the use of a fixed  $a_\phi^*(\lambda)$ , from either Bricaud *et al.* [14] or Ciotti and Bricaud [17], resulted in eigenvalues that differed more significantly from GIOP-DC, particularly in  $a_\phi(\lambda)$ . This less dynamic approach to assigning eigenvectors cannot as efficiently represent all water types at all times and, as such, these two runs appear as outliers with regard to standard deviations in the Taylor and Target diagrams (Figs. 5–7; blue circle and green square). However, while these two fixed  $a_\phi^*(\lambda)$  runs returned somewhat elevated  $\Delta b_{bp}$ ,  $\Delta a$ , and  $\Delta a_{dg}$  relative to GIOP-DC, they returned improved  $\Delta a_\phi$ . The choice of  $S_{dg}$  appears to be the most critical within the context of this experiment, at least with regard to separating total absorption into its algal and nonalgal components [38]. The three changes to  $S_{dg}$  produced the three most significant departures from GIOP-DC in retrieved eigenvalues. Reducing  $S_{dg}$  to  $0.012 \text{ nm}^{-1}$  produced the highest  $\Delta b_{bp}$  and  $\Delta a$ . This run appears as an outlier with regard to standard deviation and correlation in the Taylor and Target diagrams (Figs. 5–7; blue cross). In contrast, elevating  $S_{dg}$  to  $0.024 \text{ nm}^{-1}$  produced the highest  $\Delta a_{dg}$  and  $\Delta a_\phi$ . Dynamically selecting  $S_{dg}$  via Lee *et al.* [7] produced an equivalent  $\Delta IOP$  to GIOP-DC and the lowest overall  $\Delta a_{dg}$  of all the runs, emphasizing a potential benefit from dynamically assigning  $S_{dg}$ .

Choices made for the inversion itself also impacted the retrieved eigenvalues, a discussion of which infrequently appears in previous literature (Table 5) [9,39]. Using a linear matrix inversion method in lieu of the nonlinear LM optimization resulted in departures from GIOP-DC of only 2%–7% (MPD), with comparable  $\Delta IOPs$ . Excluding  $R_{rs}(670)$  from the inversion (that is, using only  $R_{rs}$  from 400 to 600 nm) resulted in negligible differences from GIOP-DC and similar  $\Delta IOP$ . Intuitively, we expected larger departures given the significance of red  $R_{rs}$  for SAA applications in turbid waters [1,40]. This analysis employed a synthesized data set, however, which does not comprehensively represent highly turbid, highly scattering environments. As will be explored

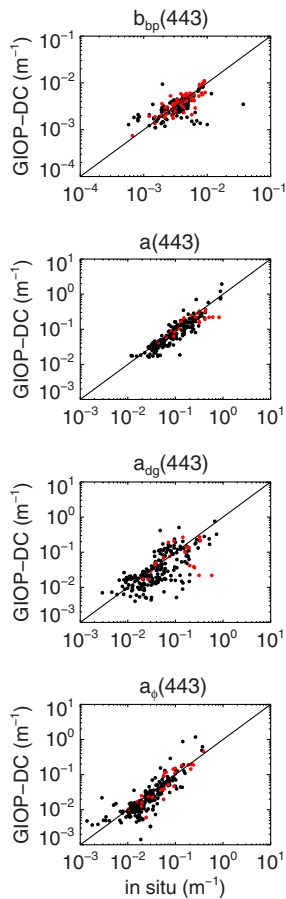


Fig. 2. Comparison of GIOP-DC and ground truth (*in situ*) IOPs at 443 nm from SeaWiFS (black) and MODISA (red). See Table 3 for accompanying statistic.

in detail later, an alternate choice of  $G(\lambda)$  [22] produced eigenvalues that deviated more significantly from GIOP-DC [6%–14% (MPD)] and resulted in elevated  $\Delta$ IOP.

#### 4. Discussion

We developed GIOP to provide a software environment for developing and applying SAAs, not to introduce a novel SAA. While GIOP-DC (one configuration instance of GIOP) provides a viable SAA for use in global applications, it remains imperfect, and we recommend that this “default configuration” evolve over time as the community develops novel methods and achieves advanced insights into bio-optical modeling and  $R_{rs}(\lambda)$ -IOP closure. We initiated this exercise to illustrate that differences in most SAAs are limited to only the selection of eigenvectors and the inversion or optimization approach, and that a consolidated software framework such as GIOP provides a simple mechanism for moving ocean color science forward by allowing controlled SAA evaluation, regional tuning, dynamic eigenvector parameterization based on optical water types (OWTs), and ensemble inversion modeling. While GIOP-DC itself performs decently for the data sets under investigation (Tables 2 and 3) and comparably to other established algorithms [41] (see Appendix A for a quantitative discussion), we cannot expect reliability at all times in all water types, as is true for all globally parameterized SAAs. We focus the following discussion on lessons learned from its development and evaluation with the goal of recommending (outlining) directions for subsequent community research.

Table 3. Regression Statistics for GIOP-DC Using the SeaWiFS and MODISA Match-Up Data Sets

	SeaWiFS					MODISA					
	<i>N</i>	$r^2$	Slope (SE)	Ratio	MPD	<i>N</i>	$r^2$	Slope (SE)	Ratio	MPD	
$b_{bp}$	412	123	0.30	0.77 (0.07)	0.93	25.2	56	0.69	0.92 (0.07)	1.00	13.2
	443	123	0.31	0.79 (0.07)	0.92	25.0	56	0.69	0.95 (0.08)	1.02	17.2
	490	123	0.32	0.82 (0.07)	0.90	24.2	56	0.69	1.00 (0.08)	0.99	18.2
	555 <sup>a</sup>	123	0.32	0.84 (0.07)	0.87	25.2	56	0.68	1.05 (0.09)	0.99	18.8
	670 <sup>b</sup>	123	0.31	0.87 (0.07)	0.83	28.5	56	0.63	1.06 (0.09)	1.04	23.0
$a$	412	192	0.74	1.12 (0.04)	0.87	30.9	21	0.45	0.76 (0.14)	0.89	36.9
	443	192	0.81	1.07 (0.03)	0.81	25.5	21	0.73	0.77 (0.10)	0.88	16.9
	490	192	0.80	1.01 (0.03)	0.76	29.3	21	0.84	0.79 (0.07)	0.79	21.1
	555 <sup>a</sup>	192	0.67	1.03 (0.05)	0.68	42.3	21	0.86	0.74 (0.07)	0.75	28.9
	670 <sup>b</sup>	180	0.69	1.19 (0.05)	0.87	45.0	17	0.47	0.91 (0.19)	1.88	87.8
$a_{dg}$	412	192	0.51	1.12 (0.06)	0.86	45.7	20	0.07	0.96 (0.27)	0.88	40.2
	443	192	0.51	1.08 (0.06)	0.78	49.8	20	0.09	0.96 (0.27)	0.81	34.8
	490	192	0.48	1.01 (0.06)	0.64	54.2	20	0.11	0.96 (0.26)	0.68	41.8
	555 <sup>a</sup>	191	0.42	0.93 (0.06)	0.50	62.5	20	0.12	0.96 (0.26)	0.46	55.3
	670 <sup>b</sup>	183	0.44	0.82 (0.05)	0.34	72.0	20	0.13	0.98 (0.26)	0.32	67.7
$a_{\phi}$	412	195	0.72	1.17 (0.05)	0.73	35.5	25	0.85	1.14 (0.09)	0.91	22.5
	443	197	0.68	1.14 (0.05)	0.80	31.5	25	0.82	1.20 (0.11)	0.90	31.0
	490	197	0.68	1.12 (0.05)	0.86	30.1	25	0.81	1.13 (0.11)	0.93	33.2
	555 <sup>a</sup>	186	0.71	1.17 (0.05)	0.95	44.8	25	0.82	1.14 (0.10)	0.97	36.6
	670 <sup>b</sup>	195	0.73	1.05 (0.04)	1.11	48.6	24	0.81	1.24 (0.12)	1.35	49.3

<sup>a</sup>indicates that the wavelength is 547 nm for MODISA.

<sup>b</sup>indicates that the wavelength is 667 nm for MODISA.



Table 4. Delta Statistics for Various Trophic Levels<sup>a</sup>

All								
	NOMAD				IOCCG			
	$N_{\%}$	$N_{total}$	Med	SIQR	$N_{\%}$	$N_{total}$	Med	SIQR
$\Delta R_{rs}$	90	964	1.68	1.05	90	500	1.04	0.49
$\Delta b_{bp}$	97	223	26.94	15.64	87	500	8.52	6.80
$\Delta a$	88	735	27.75	17.28	87	500	8.56	5.90
$\Delta a_{dg}$	90	735	51.02	28.38	87	500	27.25	20.19
$\Delta a_{\phi}$	87	780	29.32	19.17	84	500	35.83	23.25
Oligotrophic								
	NOMAD				IOCCG			
	$N_{\%}$	$N_{total}$	Med	SIQR	$N_{\%}$	$N_{total}$	Med	SIQR
$\Delta R_{rs}$	98	51	1.14	1.38	100	100	0.59	0.32
$\Delta b_{bp}$	100	1	93.27	NA	100	100	5.61	3.00
$\Delta a$	94	37	53.56	28.94	100	100	6.26	2.60
$\Delta a_{dg}$	97	37	99.83	36.75	100	100	28.38	16.06
$\Delta a_{\phi}$	93	43	20.67	11.75	100	100	22.81	11.42
Mesotrophic								
	NOMAD				IOCCG			
	$N_{\%}$	$N_{total}$	Med	SIQR	$N_{\%}$	$N_{total}$	Med	SIQR
$\Delta R_{rs}$	90	477	1.56	0.76	100	125	1.22	0.37
$\Delta b_{bp}$	97	127	33.65	15.79	100	125	6.51	3.84
$\Delta a$	93	326	25.71	16.23	100	125	6.59	3.22
$\Delta a_{dg}$	96	326	46.53	28.81	100	125	30.72	22.95
$\Delta a_{\phi}$	94	353	23.93	13.22	100	125	31.44	19.20
Eutrophic								
	NOMAD				IOCCG			
	$N_{\%}$	$N_{total}$	Med	SIQR	$N_{\%}$	$N_{total}$	Med	SIQR
$\Delta R_{rs}$	88	436	1.91	1.51	82	275	1.29	0.70
$\Delta b_{bp}$	96	95	19.90	12.06	77	275	13.31	11.04
$\Delta a$	82	372	28.25	17.31	76	275	14.56	8.77
$\Delta a_{dg}$	84	372	48.47	25.94	77	275	25.00	21.75
$\Delta a_{\phi}$	80	384	40.27	30.36	71	275	54.49	34.19

<sup>a</sup> $N_{\%}$  provides the percentage of valid retrievals returned by GIOP-DC calculated as  $100\% * N_{valid}/N_{total}$ , where  $N_{valid}$  is the number of valid retrievals (not shown).  $N_{total}$ , Med and SIQR provide the distribution sample size, median and semi-interquartile range, respectively.

GIOP provides a resource for better understanding the sensitivity of an SAA to its assigned eigenvectors. Our sensitivity analyses permitted generation of a preliminary hierarchy of eigenvector significance, which can be used to focus future research on the parameterizations of most importance. This approach cannot effectively identify an optimal set of eigenvectors and methods, as we sequentially varied just a single parameter and combinations of certain eigenvectors mutually compensate for their respective errors. Spectral shapes for  $b_{bp}(\lambda)$ ,  $a_{dg}(\lambda)$ , and  $a_{\phi}(\lambda)$  decrease from 440 to 660 nm, for example, and therefore transfer variances when their shapes change within realistic bounds. We also expect  $a_{dg}(\lambda)$  and  $a_{\phi}(\lambda)$  to covary under certain environmental conditions, such as the open ocean where the detrital pool includes derivative products of phytoplankton, such as cells walls and cytoplasm. That said, our approach identifies the eigenvectors and methods with the

largest impact on eigenvalue retrievals, and thus, reveals those requiring additional attention by the community. If we assume that the goal of any SAA is the best simultaneous retrieval of  $b_{bp}(\lambda)$ ,  $a(\lambda)$ ,  $a_{dg}(\lambda)$ , and  $a_{\phi}(\lambda)$  (e.g., minimizing  $\Delta IOP$ ), the choice of  $S_{bp}$  appears to be the least critical (Table 5). In contrast, given the robust ability of SAAs to retrieve  $b_{bp}(\lambda)$  and  $a(\lambda)$ , and their reduced capacity to deconvolve total absorption into  $a_{dg}(\lambda)$  and  $a_{\phi}(\lambda)$ , the choice of  $S_{dg}$  and  $a_{\phi}^*(\lambda)$  appears to be highly significant. Using  $\Delta IOP$  as the performance metric, the use of dynamically evolving  $S_{dg}$  and  $a_{\phi}^*(\lambda)$  produced superior results to the use of static eigenvectors (Table 5). Despite this, spectral dependencies in the retrieved eigenvalues revealed in the match-up results (Tables 2 and 3) indicates compensation for imperfect eigenvectors. Elevated  $\Delta a_{dg}$  and  $\Delta a_{\phi}$  relative to  $\Delta a$  highlights the challenge of using SAA approaches in isolation to provide information on phytoplankton

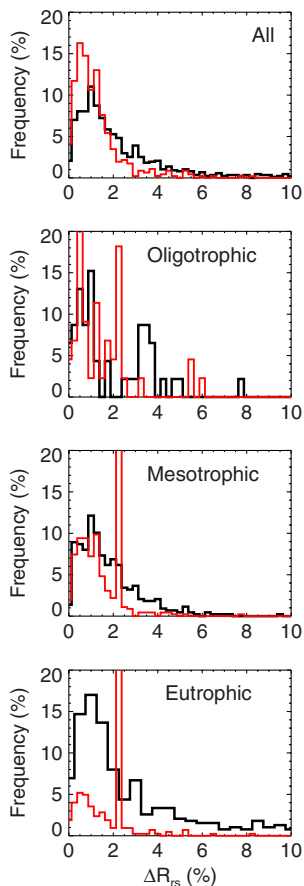


Fig. 3. Frequency distributions of  $\Delta R_{rs}$  from NOMAD (black) and the IOCCG data set (red) for the all available data (“All”) and subsetted into three trophic levels. The main text provides definitions for the oligo-, meso-, and eutrophic subsets. See Table 4 for accompanying statistics.

community structure [for example, when multiple  $a_{\phi}^*(\lambda)$  are employed to retrieve eigenvalues for multiple phytoplankton groups]. Each successive decomposition of an IOP into component IOPs adds variability and ambiguity.

The selection of  $G(\lambda)$ , and  $R_{rs}(\lambda)$ -IOP closure in general, merits additional consideration by the research community. Ambiguity in the retrieval of IOPs from  $R_{rs}(\lambda)$  remains a known issue (that is beyond the scope of this work) and combinations of successfully retrieved eigenvalues may not make geophysical sense, despite numerical closure in the inversion [42,43]. That aside, the choice of  $G(\lambda)$  from Gordon *et al.* [21] versus Morel *et al.* [22] resulted in eigenvalues that differed by 6%–14% (Table 5). An unexpected relationship between  $G(\lambda)$ ,  $a_{\phi}(\lambda)$ , and  $C_a$  emerged when evaluating these results (Fig. 8). The use of the Gordon quadratic expression [ $G_1(\lambda) = 0.0949$  and  $G_2(\lambda) = 0.0794$ ] resulted in  $a_{\phi}(443)$  with little bias across the full dynamic range of NOMAD stations (see the red best-fit line with a slope near unity in panel B), but  $C_a$  with a clear trophic bias (see the red best-fit line with a slope less than unity in panel A). In contrast, the look-up-table (LUT) approach of Morel [ $G_1(\lambda)$  from the LUT and  $G_2(\lambda) = 0$ ]

produced  $C_a$  with little bias (panel C), but  $a_{\phi}(443)$  with a definitive slope greater than unity (panel D). Such variability in derived optical and biogeochemical products calls into question the ability of existing approaches to universally estimate  $G(\lambda)$  and relate  $a_{\phi}(\lambda)$  to  $C_a$  within the SAA paradigm [39,44]. With regard to  $G(\lambda)$ , the authors acknowledge the limitations in their approaches. The quadratic coefficients from Gordon are valid for open ocean conditions and solar zenith angles  $\geq 20^\circ$ . The LUTs from Morel are valid for conditions where only phytoplankton and their derivative products shape the marine light field. Furthermore, navigating these LUTs requires an estimate of  $C_a$ , a remotely sensed data product known to be valid in the open ocean, but less so in coastal or optically complex conditions. Several alternative methods for estimating  $G(\lambda)$  in optically complex environments now exist [7,23]. While not considered in this study, GIOP provides a framework for their systematic comparison in subsequent studies. With regard to estimating phytoplankton biomass, a paucity of comprehensive, global data sets with which to obtain robust relationships confounds our ability to universally toggle between  $a_{\phi}(\lambda)$  and  $C_a$  in this paradigm. As  $C_a$  and  $a_{\phi}(443)$  remain proportionally constant in GIOP-DC [ $a_{\phi}(443) = 0.055 C_a$ ], we expect much of the variability shown in Fig. 8 to stem from variability in the *in situ* measurements. However, repeating this analysis without normalizing  $a_{\phi}^*(443)$  to  $0.055 \text{ m}^2 \text{ mg}^{-1}$  further amplified the disconnect between  $G(\lambda)$ ,  $a_{\phi}(\lambda)$ , and  $C_a$ , particularly for  $C_a > 1 \text{ mg m}^3$  (results not shown). Without evolving to more sophisticated approaches (explored below), an end-user may have to decide between tuning an SAA to optics [ $a_{\phi}(\lambda)$ ; our choice] or biogeochemistry ( $C_a$ ) (Fig. 8), both of which have merit depending on the science question(s) under consideration.

Clearly, the use of fixed eigenvectors within an SAA (or dynamic eigenvectors developed using limited data sets) cannot capture their natural variation across different optical and biogeochemical conditions. Optical properties of the oceans vary over orders of magnitude globally and remain neither spatially uniform nor constant in time [45–49]. Novel methods now exist to constrain regional variability in optical parameters, several of which port easily into the GIOP framework. OWT classification approaches provide one avenue for capturing spatial and temporal variations in optical parameters from measured  $R_{rs}(\lambda)$  [45,49]. The approach of Moore *et al.* [45], for example, assigns individual satellite pixels to specific OWTs and subsequently selects OWT-specific eigenvectors or algorithms. GIOP provides two paths to estimate the final eigenvalues for a given pixel: (1) the selected parameterizations are blended using OWT-weighted fuzzy membership functions and used to construct a single executable SAA instance; or (2) the selected parameterizations are used to construct multiple executable SAA instances and the resulting eigenvalues are blended

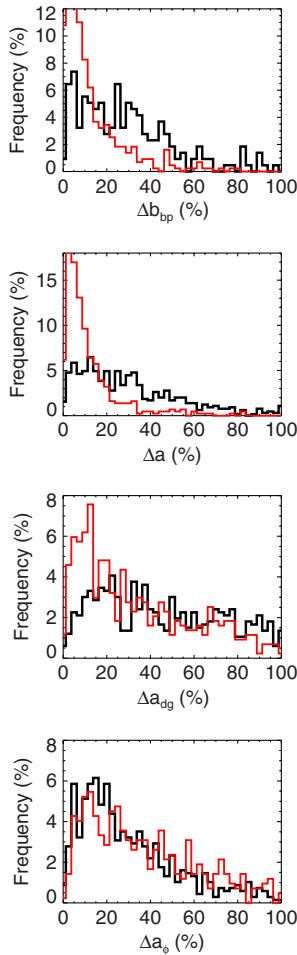


Fig. 4. Frequency distributions of  $\Delta$ IOP from NOMAD (black) and the IOCCG data set (red) for all available data. See Table 4 for accompanying statistics.

using OWT-weighted fuzzy membership functions. Ensemble (or, bootstrapping) approaches provide another vehicle to allow for spatial and temporal variations in optical parameters [9,39]. The approach

of Wang *et al.* [9], for example, successively iterates on a single satellite pixel using ranges of eigenvectors (e.g.,  $S_{dg}$  from 0.01 to 0.02  $\text{nm}^{-1}$ , and  $S_{bp}$  from  $-2$  to  $0$  at assigned intervals) and outputs the median eigenvalues (and their standard deviations) for all valid retrievals. GIOP provides a mechanism to execute the successive iterations. Both series of methods remain unconstrained by geographic boundaries and can be easily updated to accommodate additional OWT-specific parameterizations as additional *in situ* data become available.

We must consider the data sets employed in this investigation to fully interpret our results. The elevated ( $>1$ ) and reduced ( $<1$ ) ratios of  $b_{bp}(\lambda)$  and  $a(\lambda)$  suggest incomplete  $R_{rs}(\lambda)$ -IOP closure in NOMAD (Table 2). The  $R_{rs}(\lambda)$ ,  $b_{bp}(\lambda)$ , and  $a(\lambda)$  were collected using different instruments and methods, and uncertainties associated with varied measurements and methodologies are neither negligible nor spectrally independent [32]. The divergent ratios of  $a_{dg}(\lambda)$  and  $a_{\phi}(\lambda)$  with increasing wavelength likely result from low  $R_{rs}(\lambda)$  and absorption signals at long wavelengths and imperfect parameterization of  $S_{dg}$  and  $S_{bp}$ , both of which control the rate of decline of their eigenvectors. None of the data sets we considered perfectly encapsulate all bio-optical conditions at all times, nor do they represent the relative global distribution of such conditions. NOMAD remains unevenly distributed among trophic levels, showing dominance in meso- and eutrophic waters, with spatial and temporal biases [32]. The SeaWiFS and MODISA match-ups results are equally biased. The IOCCG dataset achieves optical closure as it was synthesized using a forward model with simple optical relationships (e.g., Raman and fluorescence effects were not included), but suffers similarly in how it numerically represents the relative spatial distribution of coincident bio-optical properties. Likewise, no single validation method or statistic adequately captures the full performance of an SAA. In practice, the evaluation of any algorithm

Table 5. Delta Statistics for the Sensitivity Analyses\*

Run	$N$	MPD				Median				
		$b_{bp}$	$a$	$a_{dg}$	$a_{\phi}$	$\Delta R_{rs}$	$\Delta b_{bp}$	$\Delta a$	$\Delta a_{dg}$	$\Delta a_{\phi}$
GIOP-DC	437	NA	NA	NA	NA	1.04	8.52	8.56	27.25	35.83
$S_{bp} - 33\%$	440	5.19	5.17	7.58	2.98	0.99	11.23	11.70	32.14	34.69
$S_{bp} + 33\%$	436	5.65	5.70	8.82	2.90	1.14	11.40	10.70	23.51	39.12
$S_{dg} - 33\%$	448	18.96	33.44	101.73	46.59	1.61	16.27	19.08	32.94	31.95
$S_{dg} + 33\%$	399	3.77	8.41	40.10	32.92	1.23	9.44	8.95	79.90	59.32
$S_{dg}$ from [7]	439	3.20	5.33	20.40	14.58	1.10	8.65	9.80	22.25	34.42
$C_a - 33\%$ in [14]	419	2.02	2.92	1.48	7.25	1.19	8.79	8.83	28.62	31.10
$C_a + 33\%$ in [14]	437	1.56	2.28	1.14	5.90	1.10	8.12	9.17	26.79	40.09
Fixed $C_a$ in [14]	369	4.57	7.89	2.60	21.68	1.46	11.30	11.53	30.97	26.70
$\alpha_{\phi}^*$ from [17]	357	8.33	12.72	7.04	22.23	1.20	14.26	16.75	38.30	23.13
$G$ from [22]	422	9.99	6.15	7.49	14.12	1.16	11.50	13.64	37.49	36.24
Matrix inversion	475	4.60	3.68	2.24	7.41	1.73	9.15	9.43	24.79	36.82
$400 \leq \lambda \leq 600$ nm	424	0.23	0.21	0.08	0.38	0.92	8.76	8.78	31.94	36.55

\* $N$  is the sample size. MPD is the average spectral median percent difference between GIOP-DC and each alternate run, as calculated in Tables 2 and 3. Medians of the  $\Delta$ IOP frequency distributions are also presented, as presented in Table 4.

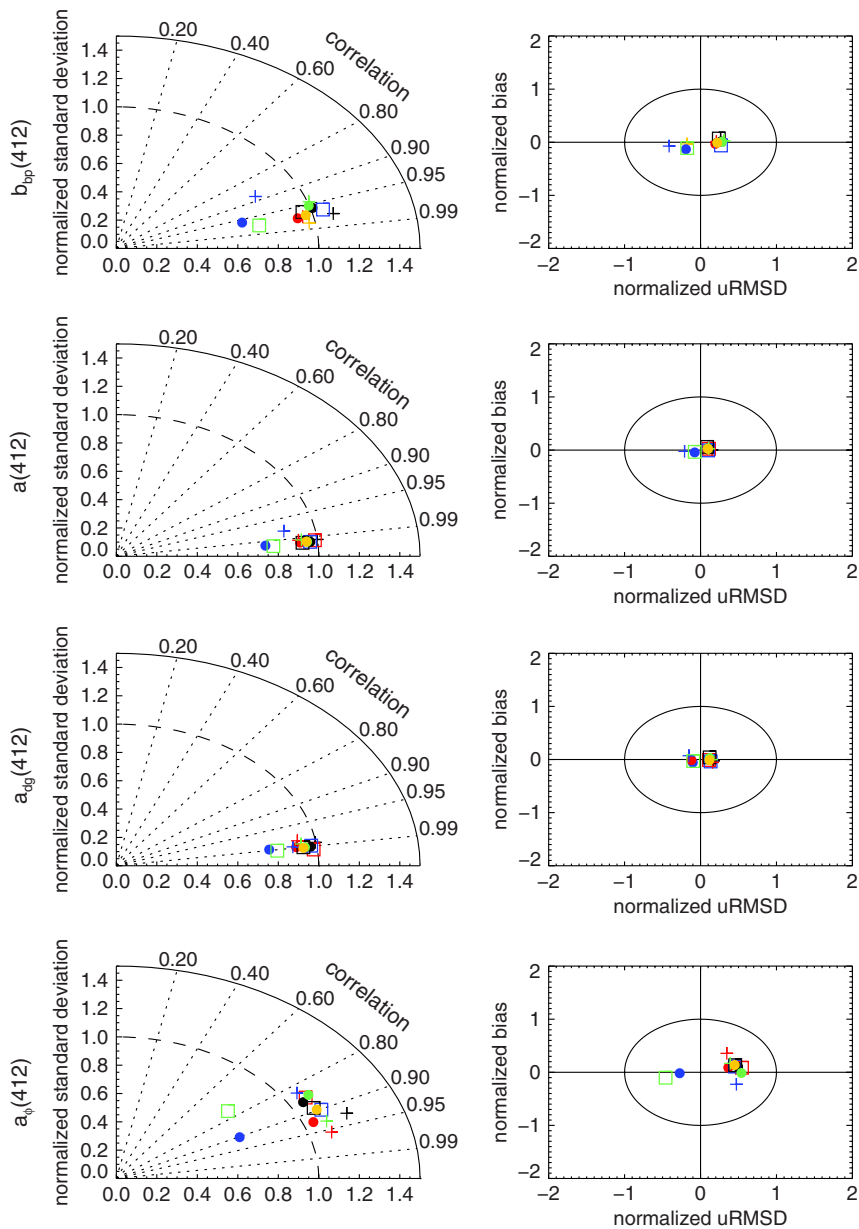


Fig. 5. Taylor and Target diagrams for IOPs at 412 nm from the IOCCG data set for the 12 alternate parameterizations of GIOP compared to GIOP-DC.  $uRMSD$  is the unbiased root mean square difference. Symbols indicate the following: blue cross =  $S_{dg} - 33\%$  ( $= 0.012 \text{ nm}^{-1}$ ); red cross =  $S_{dg} + 33\%$  ( $= 0.024 \text{ nm}^{-1}$ ); green circle =  $S_{dg}$  dynamically calculated using Lee *et al.* [7]; blue square =  $S_{bp}$  from Lee *et al.* [7]  $-33\%$ ; black square =  $S_{bp}$  from Lee *et al.* [7]  $+33\%$ ; red circle = OC-derived  $C_a - 33\%$  prior to input into Bricaud *et al.* [14]; black circle = OC-derived  $C_a + 33\%$  prior to input into Bricaud *et al.* [14]; green square =  $a_{\phi}^*(\lambda)$  from Bricaud *et al.* [14] with  $C_a$  fixed at  $0.18 \text{ mg m}^{-3}$ ; blue circle =  $a_{\phi}^*(\lambda)$  from Ciotti and Bricaud [17] with a size fraction of 0.5; black cross =  $G(\lambda)$  from Morel *et al.* [22]; orange cross = optimization using linear matrix inversion; and green cross = optimization considering only  $400 \leq \lambda \leq 600 \text{ nm}$ .

must consider the science question(s) to be addressed, which subsequently leads to the definition of validation metrics. Inevitably, trade-offs emerge, such as: (1) the value in retrieving IOPs highly accurately at a single wavelength and poorly at other wavelengths versus moderately accurately at all wavelengths; (2) the value of retrieving IOPs with high accuracy and poor spatial and temporal coverage versus moderately accurately with high spatial and temporal coverage; and (3) the value of retrieving some products with high accuracy [say,  $b_{bp}(\lambda)$

and others with less accuracy [say,  $a_{dg}(\lambda)$ ] versus all products [say, both  $b_{bp}(\lambda)$  and  $a_{dg}(\lambda)$ ] with moderate accuracy. Choices made in performance metrics ultimately affect the use and interpretation of cumulative statistics. Interpreting  $\Delta R_{rs}$ , for example, provides little value for SAAs designed to accurately retrieve IOPs at single wavelengths at the expense of quality at other wavelengths (Fig. 9).

Our choices in cost function for the inversion and goodness-of-fit metrics for the validation effort affect

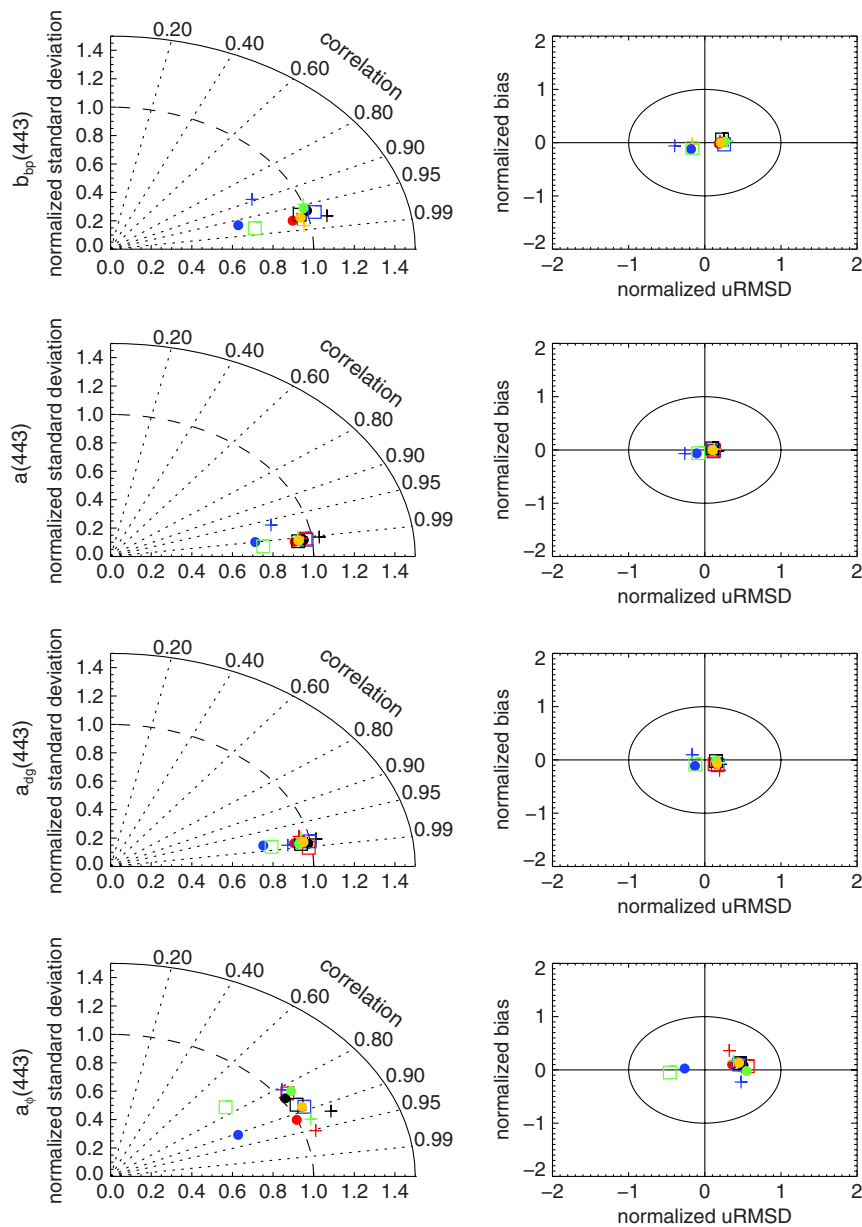


Fig. 6. As in Fig. 5, but for IOPs at 443 nm.

full interpretation of our results. The  $\chi^2$  cost function [Eq. (11)] uses absolute values and considers wavelengths from 400 to 700 nm. The  $\Delta R_{rs}$  [Eq. (10)] and  $\Delta IOP$  [Eq. (13)] metrics report relative values (%) and consider only wavelengths from 400 to 600 nm. We knowingly adopted this inconsistency for this work. In red wavelengths (>600 nm), open ocean  $R_{rs}(\lambda)$  and IOPs can be sufficiently low, such that relative (fractional) differences are large, but absolute differences are small. Conversely, in blue wavelengths, relative differences are small, while absolute differences are large. We recommend that the combined consideration of absolute and relative spectral uncertainties be adopted in future studies. Criteria for the cost function and  $\Delta R_{rs}$ , for example, could be based on maxima of assigned absolute values and relative fractions of  $R_{rs}(\lambda)$  for all wavelengths. Doing

so would not only ensure consistency between the cost function and the validation goodness-of-fit metrics, but also eliminate the need to constrain the latter to 400–600 nm.

The choice of input uncertainties for  $R_{rs}(\lambda)$  affects both SAA performance and the derived uncertainties on IOP retrievals. Reliable uncertainties for satellite-derived  $R_{rs}(\lambda)$  remain difficult to accurately determine. The use of signal-to-noise ratio (SNR) estimates for the satellite instrument and of statistical measures based on agreement between satellite retrievals and *in situ* measurements have both been explored [50–52]. However, these approaches are restricted to the assignment of a single, global uncertainty estimate to each wavelength (even the latter approach, due to the limited availability and geographic distribution of such match-ups). In practice,

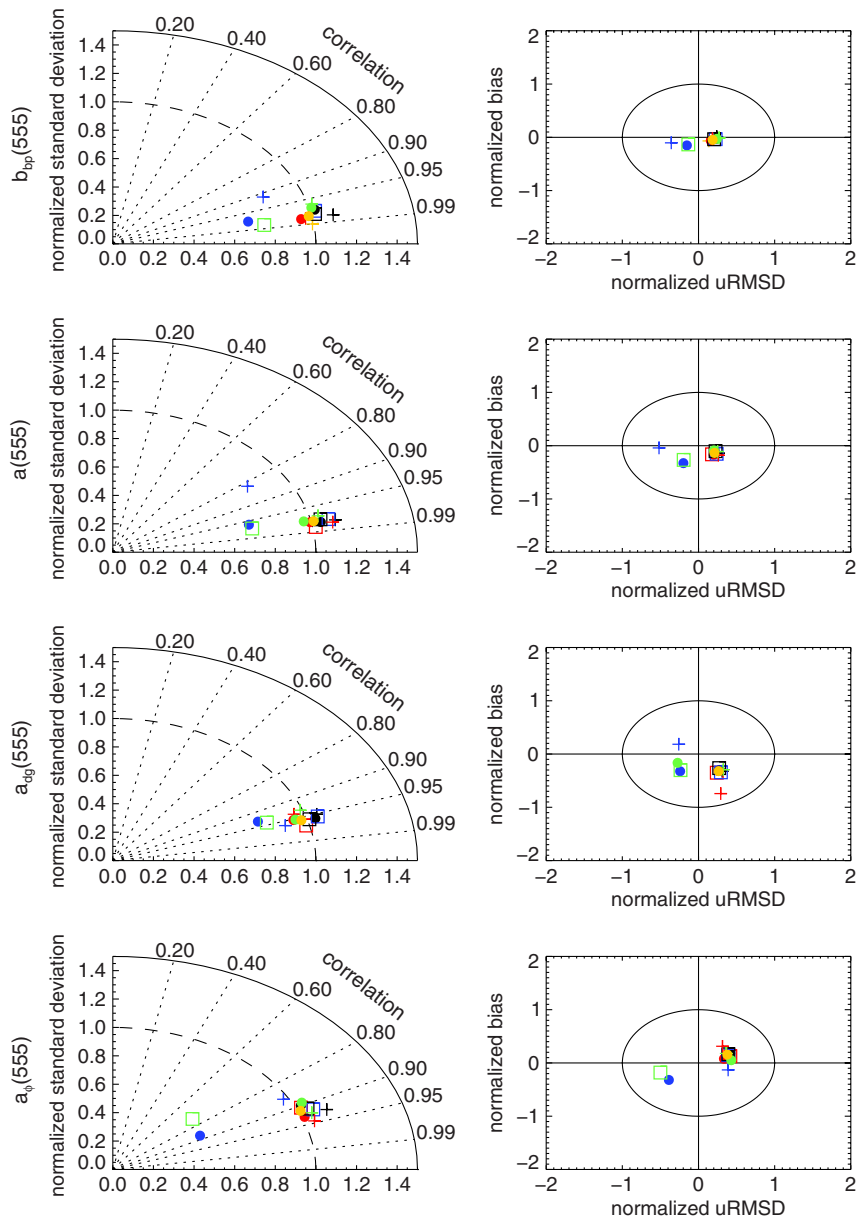


Fig. 7. As in Fig. 5, but for IOPs at 555 nm.

errors in the atmospheric correction process and changes in the instrument radiometric performance over time complicate uncertainties in satellite  $R_{rs}(\lambda)$ . Remote measurements collected at large viewing angles (large atmospheric path lengths) or through elevated aerosol loads, where the water-leaving signal is a much smaller portion of the total observed radiance at the sensor, have higher  $R_{rs}(\lambda)$  uncertainties than those collected through shorter atmospheric paths or clear atmospheres dominated by simple Rayleigh scattering. Turbid or highly reflective waters require additional corrections to separate the atmospheric signal from the water signal, which adds additional assumptions and inherent uncertainties [53]. Similarly,  $R_{rs}(\lambda)$  retrievals obtained in the vicinity of land, clouds, or sun glint may be contaminated by stray light or atmospheric adjacency effects.

Finally, the radiometric sensitivity of a satellite instrument often degrades over time as the satellite ages, leading to decreased SNRs and increased instrumental uncertainties [54]. In practice, uncertainties for  $R_{rs}(\lambda)$  vary spatially and temporally, and this variability includes changes in both their magnitude and spectral shape.

The primary impact of uncertainties for  $R_{rs}(\lambda)$  on the IOP model optimization process, as implemented for GIOP, is to change the spectral weighting of the  $\chi^2$  minimization [Eq. (11)]. For example, if the assigned variation does not represent the true relative uncertainty at each wavelength, the minimization process will be skewed to improperly favor specific wavelengths and ignore others, and the resulting model will not necessarily reproduce the spectral dependence of  $R_{rs}(\lambda)$  to within their assigned uncertainties.

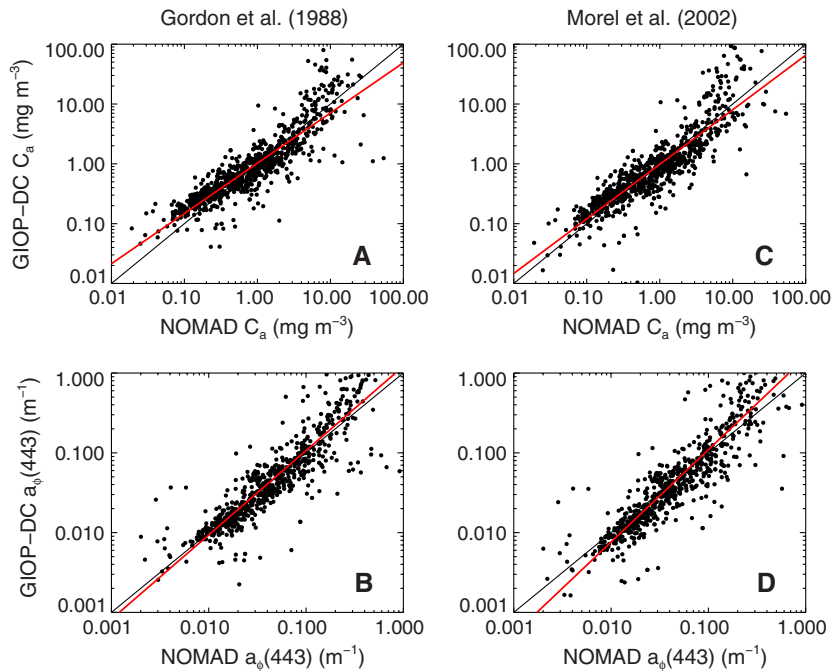


Fig. 8. Comparison of GIOP-DC and *in situ* measurements of  $C_a$  (panels A and C) and  $a_\phi(443)$  (panels B and D) from NOMAD using  $G(\lambda)$  parameterized by Gordon *et al.* [21] (panels A and B) and Morel *et al.* [22] (panels C and D). The black line indicates a 1:1 relationship. The red line indicates the best fit. The slopes of the best fit lines are 0.84, 1.06, 0.91, and 1.16 for panels A–D, respectively.

Spectrally constant error in the magnitudes of uncertainties for  $R_{rs}(\lambda)$  will have no impact on the fitting results. The derived IOP uncertainties, however, are directly proportional to the magnitude of the assigned uncertainties for  $R_{rs}(\lambda)$ .

Currently, GIOP-DC does not assign uncertainties to  $R_{rs}(\lambda)$  and uses the variance–covariance matrix to derive the IOP uncertainties. This solution effectively assumes a spectrally flat distribution for  $R_{rs}(\lambda)$  uncertainties. In practice, such an unweighted approach yields very similar IOP retrieval results as would using global mean uncertainties for  $R_{rs}(\lambda)$  from match-up analyses, as the latter tends toward relatively flat spectral dependence in the blue–green regime [51,52]. The use of the variance–covariance matrix likely understates the IOP retrieval uncertainties, since it effectively uses spectral agreement between the model and the measurements to estimate inherent variability in the  $R_{rs}(\lambda)$  measurements and, thus, ignores spectrally independent error in  $R_{rs}(\lambda)$ . The community ultimately needs a reliable method for assigning  $R_{rs}(\lambda)$  uncertainties to each remote sensing observation based on a detailed error budget of the instrument calibration and atmospheric correction process. Characterizing the  $R_{rs}(\lambda)$  match-up statistics based on viewing geometry, aerosol optical thickness, turbidity, or other relevant variables, within the limitations of the available satellite-to-*in situ* match-up database, provides an intermediate option [55,56]. Until we acquire improved knowledge of  $R_{rs}(\lambda)$  uncertainties, the assumption of global uncertainties or the use of an unweighted optimization provide viable options.

## 5. Conclusions and Future Directions

GIOP provides a consolidated software environment for developing, applying, and evaluating ocean color SAAs to retrieve marine IOPs. While our primary contribution remains the GIOP framework itself, its community-recommended default configuration,

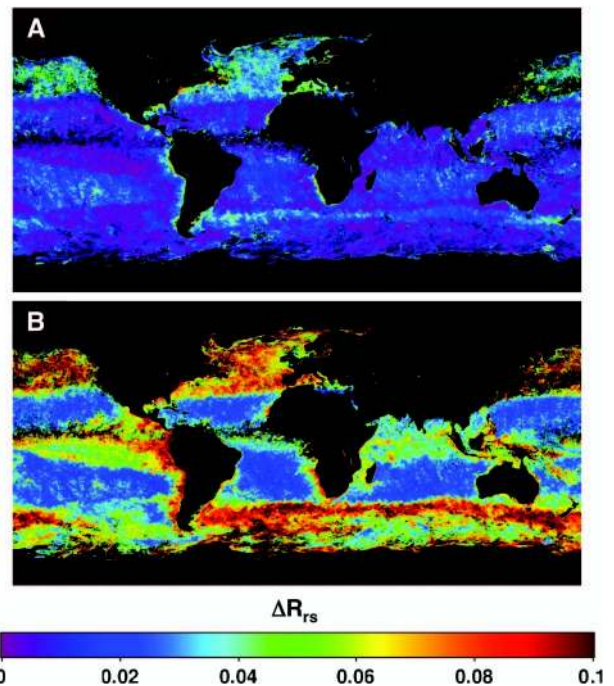


Fig. 9. MODISA  $\Delta R_{rs}$  for GIOP-DC (panel A) and GSM (run using GIOP; panel B). GIOP was applied to the monthly MODISA level-3 bin file for March 2010. Units are nondimensional (0.1 = 10%).

GIOP-DC, produces global IOPs of comparable quality to other common algorithms [41]. As GIOP can easily accommodate new eigenvectors and advanced approaches as the research community evolves, we anticipate and recommend that GIOP-DC be updated routinely. Several features recently included, or in the queue for inclusion, into GIOP include: (1) temperature and salinity dependent  $a_w(\lambda)$  and  $b_{bw}(\lambda)$  [29]; (2) alternate mathematical inversion approaches; (3) alternate eigenvector parameterizations; (4) alternate  $R_{rs}(\lambda)$ -IOP relationships [e.g.,  $G(\lambda)$  from Lee [23]]; (5) consideration of Raman inelastic scattering; (6) ensemble solution methods, such as Wang *et al.* [9] and Brando *et al.* [39]; and (7) dynamic configuration based on detected OWTs, such as Vantrepotte *et al.* [49] and Moore *et al.* [45] (provisional parameters for which were recently derived). We hope our sensitivity analyses and subsequent discussion will provide the community guidance on future directions for *in situ* data collection and algorithm refinement to support advancing SAAs (through GIOP) and their application. We also expect the GIOP framework to facilitate analyses associated with new mission planning. Its inherent ability to operate on any array of wavelengths, for example, provides a resource for identifying new channels to be added to forthcoming satellite instruments (e.g., ultraviolet bands).

## Appendix A

In practice, most common, published SAAs fall into three broad classes, hereafter referred to as spectral optimization, spectral deconvolution, and bulk inversion. In this Appendix, we briefly introduce each class with attention to the interoperability of the GIOP framework and several widely use SAAs from each class. Note, this classification includes so-called inversion algorithms [those that derive IOPs from  $R_{rs}(\lambda)$  via inverse solutions to Eqs. (2) and (3)] and does not explicitly consider empirical (statistical) or neural-network approaches.

SAAs in the spectral optimization class operate in the manner described for GIOP in Section 2B. That is, eigenvectors are predefined [e.g., for  $b_{bp}^*(\lambda)$ ,  $a_{dg}^*(\lambda)$ , and  $a_\phi^*(\lambda)$ ] and simultaneous solutions for the eigenvalues (e.g., for  $B_{bp}$ ,  $A_{dg}$ , and  $A_\phi$ ) are achieved via linear (matrix) or nonlinear (least squares) optimization of Eq. (9). The system is overdetermined if  $N_\lambda$  exceeds the number of unknowns. Examples include the SAAs described in Roesler and Perry [3], Hoge and Lyon [4], Garver and Siegel [5], Maritorena *et al.* [8], Wang *et al.* [9], and Devred *et al.* [10]. These SAAs predominantly differ in their choice of eigenvectors and inversion method and, in principle, GIOP can be configured to mimic each of them. For example, the Garver–Siegel–Maritorena (GSM) algorithm can be executed within the GIOP framework by assigning user-defined  $S_{bp}$ ,  $S_{dg}$ , and  $a_\phi^*(\lambda)$  from Maritorena *et al.* [8],  $G(\lambda)$  from Gordon *et al.* [21], and LM optimization. IOPs derived using GSM and GIOP with a GSM-like configuration compared

extremely well for the NOMAD, IOCCG, and matchup data sets (results not shown). Validation results for GIOP-DC and GSM also compared favorably with MPD for GSM-derived  $b_{bp}(443)$ ,  $a(443)$ ,  $a_{dg}(443)$ , and  $a_\phi(443)$  at 22%, 27%, 29%, and 40% for NOMAD and 22%, 26%, 20%, and 52% for the IOCCG data set (Figs. 10–12; see Table 2 for equivalent GIOP-DC statistics). The nuances in quality assurance metrics, success and failure conditions, and fail-safe behavior that accompany each SAA listed above, however, may not be currently available within the GIOP framework.

SAAs in the spectral deconvolution class similarly assign eigenvectors, but operate in a step-wise fashion to determine the spectral backscattering and absorption coefficients, rather than optimizing simultaneous solutions of the eigenvalues. Examples include the SAAs described in Lee *et al.* [7] [the quasi-analytic algorithm (QAA)], Smyth *et al.* [11], and Pinkerton *et al.* [57]. Broadly speaking, SAAs in this class operate via the following steps:

- (1) Assign  $S_{bp}$ ,  $S_{dg}$ , and  $\epsilon_\phi$  [a partial eigenvector for  $a_\phi^*(\lambda)$  defined as  $a_\phi^*(412)/a_\phi^*(443)$ ].
- (2) Estimate  $b_{bp}(\lambda_0)$ , where  $\lambda_0$  is typically a green wavelength.
- (3) Calculate  $b_{bp}(\lambda)$  as the product of  $b_{bp}(\lambda_0)$  and Eq. (8) (requires  $S_{bp}$ ).

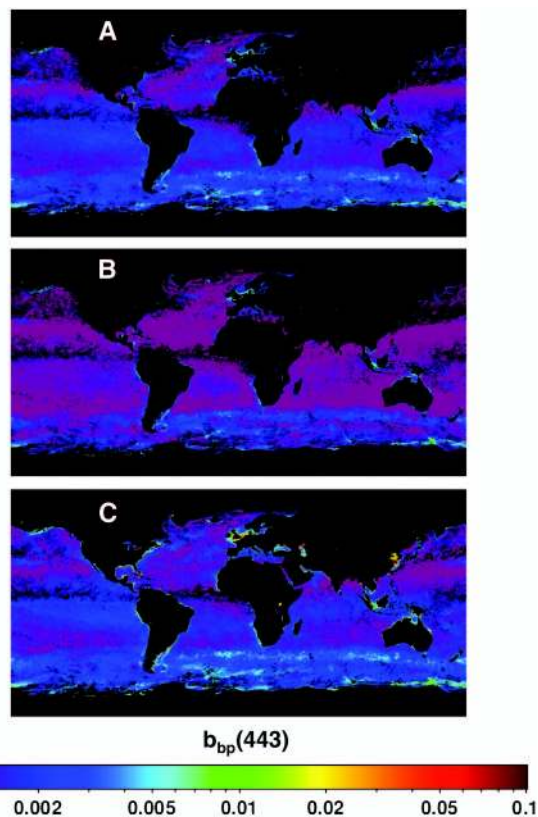


Fig. 10. MODISA  $b_{bp}(443)$  for GIOP-DC (panel A), GSM (run using GIOP; panel B), and QAA (panel C). The algorithms were applied to the monthly MODISA level-3 bin file for March 2010. Units are  $m^{-1}$ .



- (4) Calculate  $a(\lambda)$  using  $b_{bp}(\lambda)$  and Eq. (2) [requires  $G(\lambda)$ ].
- (5) Estimate  $a_{dg}(\lambda_0)$ , where  $\lambda_0$  is typically a blue wavelength (requires  $S_{dg}$  and  $\epsilon_\phi$ ).
- (6) Calculate  $a_{dg}(\lambda)$  as the product of  $a_{dg}(\lambda_0)$  and Eq. (5) (requires  $S_{dg}$ ).
- (7) Calculate  $a_\phi(\lambda)$  as  $a(\lambda) - a_w(\lambda) - a_{dg}(\lambda)$ .

Note that  $b_{bp}(\lambda_0)$  and  $a_{dg}(\lambda_0)$  are equivalent to  $B_{bp}$  and  $A_{dg}$ , respectively, in Eq. (9). These SAAs differ in their assignment of  $S_{bp}$ ,  $S_{dg}$ , and  $\epsilon_\phi$ , and in their treatment of steps 2 and 4. A complete review of the differences exceeds the scope of this paper, however, several merit mentioning. Both Smyth *et al.* [11] and Pinkerton *et al.* [57] assign constant values for  $S_{bp}$ ,  $S_{dg}$ , and  $\epsilon_\phi$ , whereas Lee *et al.* [7] dynamically estimates all three using empirical relationships based on  $R_{rs}(\lambda)$ . GIOP supports the Lee *et al.* [7] estimates of  $S_{bp}$  and  $S_{dg}$ , with the former included as part of GIOP-DC (Table 1). Both Smyth *et al.* [11] and Pinkerton *et al.* [57] adopt iterative approaches to deriving  $b_{bp}(\lambda_0)$  (step 2) and  $G(\lambda)$  (step 4). Using a LUT for  $G(\lambda)$  that is keyed on environmental geometries and absorption and scattering coefficients, both SAAs iterate until either the selected  $G(\lambda)$  or derived  $a(\lambda)$  stabilize. In contrast, Lee [7] adopts  $G(\lambda)$  from Gordon *et al.* [21] with modified coefficients and dynamically estimates  $a(\lambda_0)$  using an empirical relationship based on  $R_{rs}(\lambda)$ , which is in turn used to derive  $b_{bp}(\lambda_0)$  via rearrangement of Eqs. (2) and (3). Note that SAAs in this class can be halted at step 4

to enable testing or application of alternate approaches to decompose  $a(\lambda)$  into its component parts (e.g., [17] and [58]). At this time, GIOP does not support the step-wise deconvolution approach typical of this SAA class. However, validation results for GIOP-DC and QAA compared favorably with MPD for QAA-derived  $b_{bp}(443)$ ,  $a(443)$ ,  $a_{dg}(443)$ , and  $a_\phi(443)$  at 38%, 21%, 30%, and 28% for NOMAD and 16%, 14%, 19%, and 61% for the IOCCG data set (Figs. 10–12; see Table 2 for equivalent GIOP-DC statistics). Per their design, SAAs in this class always report  $\Delta R_{rs} = 0$ , unlike SAAs in the spectral optimization class.

SAAs in the bulk inversion class do not assign eigenvectors, that is, they do not predefine spectral shapes for the absorption or scattering coefficients. The approach introduced in Loisel and Stramski (LAS) [6] provides a widely used example. Briefly, LAS exploits a relationship between the diffuse attenuation coefficient for downwelling irradiance,  $K_d(\lambda)$  ( $m^{-1}$ ), solar zenith angle, and the absorption and scattering coefficients (e.g., [59]). As such, it requires the remote estimation of  $K_d(\lambda)$  from  $R_{rs}(\lambda)$ . LAS sequentially estimates  $b_{bp}(\lambda_i)$  and  $a(\lambda_i)$  at each wavelength  $\lambda_i$  using  $K_d(\lambda_i)$ , environmental geometries, and LUTs derived from radiative transfer (Monte Carlo) simulations. By estimating IOPs at each wavelength independently, spectral shape functions, such as  $S_{bp}$ , can be calculated dynamically and considered output products [60]. GIOP supports the use of LAS-derived  $S_{bp}$  and (tabulated)  $b_{bp}^*(\lambda)$  as

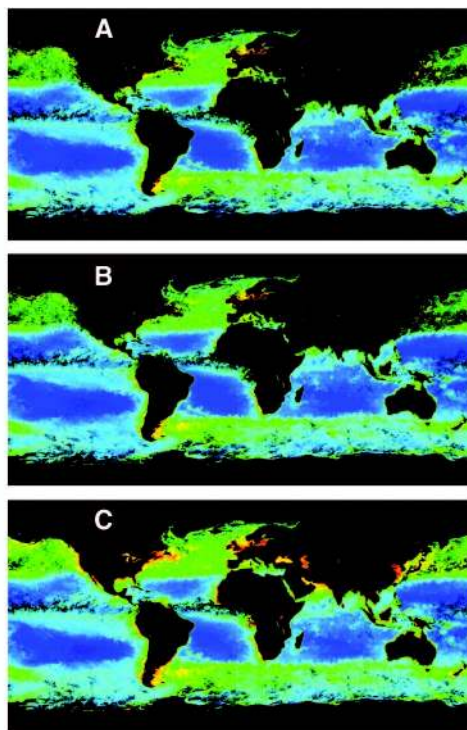


Fig. 11. As in Figure 10, but for  $a_{dg}(443)$ . Units are  $m^{-1}$ .

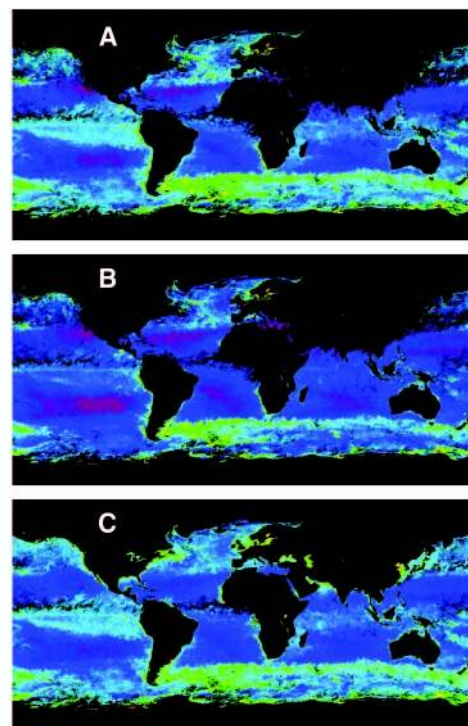


Fig. 12. As in Figure 10, but for  $a_\phi(443)$ . Units are  $m^{-1}$ .

assignable eigenvectors (Table 1). GIOP does not support, however, the spectrally independent inversion approach typical of this SAA class.

We thank Mike Behrenfeld, Paula Bontempi, Catherine Brown, Yannick Huot, Paul Lyon, Constant Mazeran, and Jill Schwarz for their helpful advice and participation in the NASA GIOP workshops. We also thank an anonymous reviewer for useful comments that improved this manuscript. Support for this work was provided through the NASA MODIS Science Team (P.J.W., B.A.F., S.W.B.) and the CSIRO Wealth from Oceans Flagship (V.B.).

## References

1. IOCCG, "Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and applications," Reports of the International Ocean-Colour Coordinating Group No. 5 (IOCCG, 2006).
2. P. J. Werdell, "Global bio-optical algorithms for ocean color satellite applications," *EOS Trans. AGU* **90**, 4 (2009).
3. C. S. Roesler and M. J. Perry, "In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance," *J. Geophys. Res.* **100**, 13279–13294 (1995).
4. 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.* **101**, 16631–16648 (1996).
5. 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.* **102**, 18607–18625 (1997).
6. 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).
7. Z.-P. Lee, K. L. Carder, and R. Arnone, "Deriving inherent optical properties from water color: a multi-band quasi-analytical algorithm for optically deep waters," *Appl. Opt.* **41**, 5755–5772 (2002).
8. S. Maritorena, D. A. Siegel, and A. Peterson, "Optimization of a semi-analytic ocean color model for global scale applications," *Appl. Opt.* **41**, 2705–2714 (2002).
9. P. Wang, E. Boss, and C. S. Roesler, "Uncertainties of inherent optical properties obtained from semi-analytical inversions of ocean color," *Appl. Opt.* **44**, 4074–4085 (2005).
10. E. Devred, S. Sathyendranath, and T. Platt, "Inversion based on a semi-analytical reflectance model," in Reports of the International Ocean-Colour Coordinating Group No. 5, Z.-P. Lee, ed. (IOCCG, 2006), pp. 87–94.
11. 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).
12. C. R. McClain, G. C. Feldman, and S. B. Hooker, "An overview of the SeaWiFS Project and strategies for producing a climate research quality global ocean bio-optical time series," *Deep Sea Res. II* **51**, 5–42 (2004).
13. B. A. Franz and P. J. Werdell, "A generalized framework for modeling of inherent optical properties in ocean remote sensing applications," in Proceedings of Ocean Optics, Anchorage, Alaska, 27 September–1 October 2010, 2010.
14. A. Bricaud, A. Morel, M. Babin, K. Allali, and H. Claustre, "Variations in light absorption by suspended particles with chlorophyll *a* concentration in oceanic (case 1) waters: analysis and implications for bio-optical models," *J. Geophys. Res.* **103**, 31033–31044 (1998).
15. A. M. Ciotti, J. J. Cullen, and M. R. Lewis, "A semi-analytical model of the influence of phytoplankton community structure on the relationship between light attenuation and ocean color," *J. Geophys. Res.* **104**, 1559–1578 (1999).
16. A. Morel and S. Maritorena, "Bio-optical properties of oceanic waters: a reappraisal," *J. Geophys. Res.* **106**, 7163–7180 (2001).
17. A. M. Ciotti and A. Bricaud, "Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil," *Limnol. Oceanogr.* **4**, 237–253 (2006).
18. A. Magnuson, L. W. Harding, Jr., M. E. Mallonee, and J. E. Adolf, "Bio-optical model for Chesapeake Bay and the middle Atlantic bight," *Estuar. Coast. Shelf. Sci.* **61**, 403–424 (2004).
19. B. A. Franz, "l2gen: the Multi-Sensor Level-1 to Level-2 Generator," [http://oceancolor.gsfc.nasa.gov/WIKI/OCSSW\(2f\)l2gen.html](http://oceancolor.gsfc.nasa.gov/WIKI/OCSSW(2f)l2gen.html) (2010).
20. K. Baith, R. Lindsay, G. Fu, and C. R. McClain, "Data analysis system developed for ocean color satellite sensors," *EOS. Trans. AGU* **82**, 202 (2001).
21. 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.* **93**, 10909–10924 (1988).
22. A. Morel, D. Antoine, and B. Gentili, "Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function," *Appl. Opt.* **41**, 6289–6306 (2002).
23. Z.-P. Lee, K. Du, K. J. Voss, G. Zibordi, B. Lubac, R. Arnone, and A. Weidemann, "An inherent-optical-property-centered approach to correct the angular effects in water-leaving radiance," *Appl. Opt.* **50**, 3155–3167 (2011).
24. C. S. Roesler, M. J. Perry, and K. L. Carder, "Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters," *Limnol. Oceanogr.* **34**, 1510–1523 (1989).
25. A. Bricaud and A. Morel, "Light attenuation and scattering by phytoplanktonic cells: a theoretical modeling," *Appl. Opt.* **25**, 571–580 (1986).
26. D. Doxaran, K. Ruddick, D. McKee, B. Gentili, D. Tailliez, M. Chami, and M. Babin, "Spectral variations of light scattering by marine particles in coastal waters, from visible to near infrared," *Limnol. Oceanogr.* **54**, 1257–1271 (2009).
27. 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).
28. 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).
29. X. Zhang, L. Hu, and M.-X. He, "Scattering by pure seawater: effect of salinity," *Opt. Express* **17**, 5698–5710 (2009).
30. S. Sugihara and M. Kishino, "An algorithm for estimating the water quality parameters from irradiance just below the sea surface," *J. Geophys. Res.* **93**, 10857–10862 (1988).
31. J. E. O'Reilly, S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. R. McClain, "Ocean color chlorophyll algorithms for SeaWiFS," *J. Geophys. Res.* **103**, 24937–24953 (1998).
32. 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 Sens. Environ.* **98**, 122–140 (2005).
33. S. W. Bailey and P. J. Werdell, "A multi-sensor approach for the on-orbit validation of ocean color satellite data products," *Remote Sens. Environ.* **102**, 12–23 (2006).
34. B. A. Franz, "NASA ocean color data reprocessing," <http://oceancolor.gsfc.nasa.gov/WIKI/OCReproc.html> (2010).
35. P. J. Werdell, S. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, G. Feldman, and C. McClain, "Unique data repository facilitates ocean color satellite validation," *EOS Trans. AGU* **84**, 377 (2003).
36. K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *J. Geophys. Res.* **106**, 7183–7192 (2001).
37. J. K. Jolliff, J. C. Kindle, I. Shulman, B. Penta, M. A. M. Friedrichs, R. Helber, and R. A. Arnone, "Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment," *J. Mar. Syst.* **76**, 64–82 (2009).

38. Z. P. Lee, R. Arnone, C. Hu, P. J. Werdell, and B. Lubac, "Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm," *Appl. Opt.* **49**, 369–381 (2010).
39. V. E. Brando, A. G. Dekker, Y. J. Park, and T. Schroeder, "Adaptive semianalytic inversion of ocean color radiometry in optically complex waters," *Appl. Opt.* **51**, 2808–2833 (2012).
40. P. J. Werdell, S. W. Bailey, B. A. Franz, L. W. Harding, Jr., G. C. Feldman, and C. R. McClain, "Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua," *Remote Sens. Environ.* **113**, 1319–1330 (2009).
41. R. J. W. Brewin, S. Sathyendranath, D. Müller, C. Brockmann, P.-Y. Deschamps, E. Devred, R. Doerffer, N. Fomferra, B. Franz, M. Grant, S. Groom, A. Horseman, C. Hu, H. Krasemann, Z.-P. Lee, S. Maritorena, F. Mélin, M. Peters, T. Platt, P. Regner, T. Smyth, F. Steinmetz, J. Swinton, J. Werdell, and G. N. White III, "The ocean colour climate change initiative: a round-robin comparison of in-water bio-optical algorithms," *Rem. Sens. Environ.* (2012) (to be published).
42. M. Sydor, R. W. Gould, R. A. Arnone, V. I. Haltrin, and W. Goode, "Uniqueness in remote sensing of the inherent optical properties of ocean water," *Appl. Opt.* **43**, 2156–2162 (2004).
43. M. Defoin-Platel and M. Chami, "How ambiguous is the inverse problem of ocean color in coastal waters?" *J. Geophys. Res.* **112**, C03004 (2007).
44. R. J. W. Brewin, E. Devred, S. Sathyendranath, S. J. Lavender, and N. J. Hardman-Mountford, "Model of phytoplankton absorption based on three size classes," *Appl. Opt.* **50**, 4535–4549 (2011).
45. T. S. Moore, J. W. Campbell, and M. D. Dowell, "A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product," *Remote Sens. Environ.* **113**, 2424–2430 (2009).
46. H. Loisel, B. Lubac, D. Dessailly, L. Duforet-Gaurier, and V. Vantrepotte, "Effect of inherent optical properties variability on the chlorophyll retrieval from ocean color remote sensing: an in situ approach," *Opt. Express* **18**, 20949–20959 (2010).
47. M. Szeto, P. J. Werdell, T. S. Moore, and J. W. Campbell, "Are the world's oceans optically different?," *J. Geophys. Res.* **116**, C00H04 (2011).
48. M. J. Sauer, C. S. Roesler, P. J. Werdell, and A. Barnard, "Under the hood of satellite empirical chlorophyll-a algorithms: revealing the dependencies of maximum band ratio algorithms on inherent optical properties," *Opt. Express* **20**, 20920–20933 (2012).
49. V. Vantrepotte, H. Loisel, D. Dessailly, and X. Mériaux, "Optical classification of constricted coastal waters," *Remote Sens. Environ.* **123**, 306–323 (2012).
50. M. L. Estapa, "Photochemical reactions of particulate organic matter," Ph.D. dissertation (University of Maine, 2011).
51. S. Maritorena, O. H. F. d'Andon, A. Mangin, and D. A. Siegel, "Merged satellite ocean color data products using a bio-optical model: characteristics, benefits, and uses," *Rem. Sens. Environ.* **114**, 1791–1804 (2010).
52. F. Mélin, G. Zibordi, and S. Djavidnia, "Merged series of normalized water leaving radiances obtained from multiple satellite missions for the Mediterranean Sea," *Adv. Space Res.* **43**, 423–437 (2009).
53. 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).
54. G. Meister, B. A. Franz, E. J. Kwiatkowska, and C. R. McClain, "Corrections to the calibration of MODIS aqua ocean color bands derived from SeaWiFS data," *IEEE Trans. Geosci. Remote Sens.* **50**, 310–319 (2012).
55. F. Mélin, G. Zibordi, and J. F. Berthon, "Assessment of satellite ocean color products at a coastal site," *Remote Sens. Environ.* **110**, 192–215 (2007).
56. G. Zibordi, F. Mélin, and J. F. Berthon, "Intra-annual variations of biases in remote sensing primary ocean color products at a coastal site," *Remote Sens. Environ.* **124**, 624–636 (2012).
57. 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," *N. Z. J. Mar. Freshwater Res.* **40**, 227–247 (2006).
58. E. Devred, S. Sathyendranath, V. Stuart, H. Maass, O. Ulloa, and T. Platt, "A two-component model of phytoplankton absorption in the open ocean: Theory and applications," *J. Geophys. Res.* **111**, C03011 (2006).
59. J. T. O. Kirk, "Dependence of relationship between inherent and apparent optical properties of water on solar altitude," *Limnol. Oceanogr.* **29**, 350–356 (1984).
60. H. Loisel, J.-M. Nicolas, A. Sciandra, D. Stramski, and A. Poteau, "Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean," *J. Geophys. Res.* **111**, C09024 (2006).