AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2014JC010604

Key Points:
- A partitioning model for light absorption coefficient of natural waters
- Nonalgal particulate and CDOM components are separated
- No restrictive assumptions about absorption coefficients are required

Supporting Information:
- Supporting Information S1

Correspondence to:
G. Zheng,
guangming.zheng@noaa.gov

Citation:
Zheng, G., D. Stramski, and P. M. DiGiacomo (2015), A model for partitioning the light absorption coefficient of natural waters into phytoplankton, nonalgal particulate, and colored dissolved organic components: A case study for the Chesapeake Bay, *J. Geophys. Res. Oceans*, *120*, 2601–2621, doi:10.1002/2014JC010604.

Received 19 NOV 2014
Accepted 22 JAN 2015
Accepted article online 27 JAN 2015
Published online 3 APR 2015

A model for partitioning the light absorption coefficient of natural waters into phytoplankton, nonalgal particulate, and colored dissolved organic components: A case study for the Chesapeake Bay

Guangming Zheng[1,2,3], Dariusz Stramski[4], and Paul M. DiGiacomo[1]

[1]NOAA/NESDIS/Center for Satellite Application and Research, College Park, Maryland, USA, [2]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA, [3]Now at Global Science and Technology Inc., Greenbelt, Maryland, USA, [4]Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA

Abstract We present a model, referred to as Generalized Stacked-Constraints Model (GSCM), for partitioning the total light absorption coefficient of natural water (with pure-water contribution subtracted), $a_{nw}(\lambda)$, into phytoplankton, $a_{ph}(\lambda)$, nonalgal particulate, $a_d(\lambda)$, and CDOM, $a_g(\lambda)$, components. The formulation of the model is based on the so-called stacked-constraints approach, which utilizes a number of inequality constraints that must be satisfied simultaneously by the model outputs of component absorption coefficients. A major advancement is that GSCM provides a capability to separate the $a_d(\lambda)$ and $a_g(\lambda)$ coefficients from each other using only weakly restrictive assumptions about the component absorption coefficients. In contrast to the common assumption of exponential spectral shape of $a_d(\lambda)$ and $a_g(\lambda)$ in previous models, in our model these two coefficients are parameterized in terms of several distinct spectral shapes. These shapes are determined from field data collected in the Chesapeake Bay with an ultimate goal to adequately account for the actual variability in spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$ in the study area. Another advancement of this model lies in its capability to account for potentially nonnegligible magnitude of $a_d(\lambda)$ in the near-infrared spectral region. Evaluation of model performance demonstrates good agreement with measurements in the Chesapeake Bay. For example, the median ratio of the model-derived to measured $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ at 443 nm is 0.913, 1.064, and 1.056, respectively. Whereas our model in its present form can be a powerful tool for regional studies in the Chesapeake Bay, the overall approach is readily adaptable to other regions or bio-optical water types.

1. Introduction

Suspended and dissolved constituents in natural waters such as phytoplankton, detritus, mineral particles, and colored dissolved organic matter (CDOM) affect the status and functioning of aquatic ecosystems. The ability to accurately quantify the presence and amounts of these materials in water is critical for numerous science questions and applications, such as in the area of aquatic ecology and biogeochemistry [e.g., *Falkowski et al*., 1998; *Moore et al*., 2004; *Coble*, 2007], and monitoring of water quality [e.g., *Hu et al*., 2004; *Schaeffer et al*., 2012]. Optical measurements that can be taken from in situ, airborne, or satellite platforms provide an efficient tool for monitoring and characterizing major categories of water constituents, with an important capability for providing information over extended temporal and spatial scales. In particular, the reflectance of a water body determined from in-water, airborne, or satellite radiometric measurements carries information about the inherent optical properties (IOPs) of water such as the total spectral light absorption, $a(\lambda)$, and backscattering, $b_b(\lambda)$, coefficients, where λ denotes the wavelength of light in vacuum. The total absorption coefficient (or equivalently the total nonwater absorption coefficient, $a_{nw}(\lambda) = a(\lambda) - a_w(\lambda)$ where $a_w(\lambda)$ is the known absorption coefficient of pure water) can in turn be partitioned into phytoplankton and nonphytoplankton components [e.g., *Roesler et al*., 1989; *Lee et al*., 2002; *Ciotti and Bricaud*, 2006; *Zheng and Stramski*, 2013a].

One major challenge for partitioning the total absorption coefficient $a(\lambda)$ [or equivalently $a_{nw}(\lambda)$] into component contributions is to separate the nonalgal particulate absorption, $a_d(\lambda)$, and CDOM absorption, $a_g(\lambda)$,

coefficients from one another. The main cause of the difficulty in separating $a_d(\lambda)$ and $a_g(\lambda)$ is related to the similarity of spectral shapes of these two coefficients, which are typically approximated by the exponential function of light wavelength [e.g., *Babin et al*., 2003]. As a result, the existing models for partitioning $a(\lambda)$ typically yield the phytoplankton absorption coefficient, $a_{ph}(\lambda)$, and the combined nonphytoplankton absorption coefficient, $a_{dg}(\lambda) = a_d(\lambda) + a_g(\lambda)$ [*Roesler et al*., 1989; *Lee et al*., 2002; *Ciotti and Bricaud*, 2006; *Zheng and Stramski*, 2013a]. There have also been a few attempts to derive separate contributions of $a_d(\lambda)$ and $a_g(\lambda)$ from partitioning of $a(\lambda)$ [*Gallegos and Neale*, 2002; *Schofield et al*., 2004; *Lin et al*., 2013]. These models have, however, several significant limitations, most notably highly restrictive assumptions about model outputs of phytoplankton and nonphytoplankton absorption coefficients. For example, for $a_{ph}(\lambda)$ the assumptions involve a single spectral shape [*Gallegos and Neale*, 2002] or a linear combination of a small number of predefined spectral shapes [*Schofield et al*., 2004].

In the most recent model of *Lin et al*. [2013], $a_{ph}(\lambda)$ is parameterized with an empirical quadratic function of $a_{ph}(\lambda_0)$ on the basis of a large set of field data, where λ_0 is a reference light wavelength chosen at 489 nm. This parameterization seriously limits the variation in the spectral shape of model-derived $a_{ph}(\lambda)$. For example, for the observed range of $a_{ph}(\lambda_0)$ values ($\leq 1\ m^{-1}$) in the large data set, the blue-to-red spectral ratio of phytoplankton absorption, $a_{ph}(443)$:$a_{ph}(670)$, is allowed to vary only between about 1.72 and 1.83. In this model, the total particulate absorption coefficient, $a_p(\lambda)$, is similarly parameterized in terms of empirical quadratic function involving the magnitude of the coefficient at a reference wavelength. In addition, each individual spectrum of $a_d(\lambda)$ and $a_g(\lambda)$ is described in terms of exponential function of λ with a single slope value (S_d and S_g, respectively) within a large spectral range, albeit variations in these slope parameters are allowed on a case-by-case basis (these assumptions are similar to those used by *Schofield et al*. [2004]). Such parameterization with a single slope value fails to account for potential changes in the slope as a function of wavelength within a given spectrum [e.g., *Loiselle et al*., 2009]. Another potential weakness of the model by *Lin et al*. [2013] is that the closure described by the equation $a_p(\lambda) = a_{ph}(\lambda) + a_d(\lambda)$ does not appear to be necessarily satisfied because two sets of solutions for $a_{ph}(\lambda)$ and $a_d(\lambda)$ are derived by the model, which are not necessarily identical. Also, whereas these investigators report the root-mean-square errors ranging generally from about 10% to 40% for model-derived $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$ based on global and regional field data sets, no systematic errors are reported. However, the graphical representation of model results indicates significant bias in $a_{ph}(\lambda)$ and $a_d(\lambda)$.

The limitations of the existing partitioning models adversely affect their performance and emphasize the need for a less restrictive and more general approach for partitioning the absorption coefficient into phytoplankton, nonalgal particulate, and CDOM components. This level of discrimination in turn provides more quantitative information for both research and applied uses, particularly in the fields of marine biogeochemistry, water quality, and broader coastal resource management. In this paper, we describe and validate such an approach for partitioning $a_{nw}(\lambda)$ into $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$. The main prerequisite for developing such an improved partitioning model is to relax restrictive assumptions about the spectral shapes of model outputs of $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$. This is because it is critical for the model to adequately account for large variations in the spectral shape of $a_{ph}(\lambda)$ which can be caused by changes in phytoplankton community composition and acclimation to environmental conditions [e.g., *Geider et al*., 1998; *Stramski et al*., 2001; *Babin et al*., 2003]. In addition, variations in the spectral shape of $a_d(\lambda)$ are possible, which can depart significantly from a simple exponential function of light wavelength and exhibit nonzero magnitude in the near-infrared (NIR), especially when the particle assemblage includes absorbing minerals [e.g., *Iturriaga and Siegel*, 1989; *Babin and Stramski*, 2004; *Bowers and Binding*, 2006; *Stramski et al*., 2007]. Two partitioning models have been developed recently on the basis of the so-called stacked-constraints approach that is able to account for large variability in the spectral shapes of model outputs of component absorption coefficients [*Zheng and Stramski*, 2013a, 2013b]. One of these models partitions $a_{nw}(\lambda)$ into $a_{ph}(\lambda)$ and $a_{dg}(\lambda)$ [*Zheng and Stramski*, 2013a] and the other partitions $a_p(\lambda)$ into $a_{ph}(\lambda)$ and $a_d(\lambda)$ [*Zheng and Stramski*, 2013b]. A key common feature of these models is that a set of appropriately defined inequality constraints is used to relax the restrictive assumptions about the component absorption coefficients and to solve the partitioning problem with greatly improved accuracy. The validation based on a large set of field data demonstrated good performance of both models. We refer to this type of model as the stacked-constraints model (SCM) and in this paper we extend this approach to the development of a model for partitioning $a_{nw}(\lambda)$ into the three separate components of $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$.

As a proof-of-concept study to demonstrate the feasibility of partitioning $a_{nw}(\lambda)$ into $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$, we selected a coastal region of significant socioeconomic importance, namely the Chesapeake Bay which is the largest estuary in the United States (U.S.). Its watershed comprises urbanized, agricultural, forested, and mixed-use regions within several highly populated states (e.g., Maryland, Virginia, Pennsylvania, and Washington, D.C.). The bay is routinely monitored and studied, meaning a wealth of supporting in situ measurements is available.

Coastal, estuarine, and inland waters are optically more complex than open ocean waters, imposing additional challenges for deriving information about different water constituents from measurements of bulk optical properties. The nonalgal particulate and CDOM constituents of water do not typically covary with phytoplankton and the spectral shapes of light absorption coefficients of these components are usually highly variable in coastal and inland aquatic environments owing to significant terrigenous as well as anthropogenic inputs to local receiving waters. Despite these complexities, the selection of an optically complex coastal environment for our study can offer specific advantages for the purpose of separating $a_d(\lambda)$ and $a_g(\lambda)$ from one another. This is because the presence of mineral particles in such waters can produce distinct spectral features or significant departures from an exponential spectral shape of $a_d(\lambda)$. These features can aid in distinguishing the $a_d(\lambda)$ spectrum from the $a_g(\lambda)$ spectrum that follows the exponential shape more closely.

Using a large set of measurements of component absorption coefficients from the Chesapeake Bay, we demonstrate and validate a partitioning model based on the stacked-constraints approach that involves only very weakly restrictive assumptions with respect to actual variability of absorption properties within the study area. Although this is a case study with a specific regional focus, the approach presented for the partitioning model can be adapted to other regions or bio-optical water types. Thus, this capability can provide an improved basis for pursuing and supporting a variety of scientific and applied-oriented problems in various aquatic environments on the basis of optical measurements from both in situ and remote-sensing platforms.

2. Field Data of Absorption Coefficients

To develop the partitioning model, field data of absorption coefficients are used to: (i) characterize spectral shapes of absorption coefficients, (ii) determine the inequality constraints, and (iii) evaluate the performance of the model. For these purposes, we selected four sets of absorption data collected at near-surface depths (0–3 m) in the Chesapeake Bay. Our selection is limited to near-surface data to ensure the suitability of our model to remote-sensing applications. Measured variables in the selected data sets include the spectral absorption coefficients of nonalgal particles $a_d(\lambda)$ (i.e., depigmented particulate component), CDOM $a_g(\lambda)$, and phytoplankton $a_{ph}(\lambda)$. These coefficients were measured at high spectral resolution (1 or 2 nm) using methods summarized below in section 2.1. To ensure uniform spectral resolution of 1 nm for all data, linear interpolation was applied to data with original spectral resolution of 2 nm.

The entire data set includes measurements taken at 451 stations within the Chesapeake Bay (Figure 1). The data were collected as part of four projects including the Land Margin Ecosystem Response—Trophic Interactions in Estuarine Systems (LMER-TIES, 1996–2000) [*Harding and Magnuson*, 2003], Bio-complexity (BIOCOMP, 2001–2004), Atlantic Coast Estuaries—Indicators Consortium (ACE-INC, 2002–2003) [*Harding et al*., 2005], and GEO-stationary Coastal and Air Pollution Events (GEO-CAPE) Chesapeake Bay oceanographic campaign in July 2011 [*Le et al*., 2013].

2.1. Measurements and Data Processing

Methods to obtain the data of absorption coefficients are consistent with the protocols described by *Harding and Magnuson* [2003] and *Harding et al*. [2005] for the LMER-TIES, BIOCOMP, and ACE-INC data sets, and by *Mannino et al*. [2008] and *Le et al*. [2013] for the GEO-CAPE data set. Here we provide a brief summary of the methods. Spectra of total particulate absorption coefficient, $a_p(\lambda)$, were determined with a quantitative filter pad method by spectrophotometric measurements of the transmittance of light through a GF/F filter containing particles, i.e., T-technique [e.g., *Mitchell*, 1990; *Mitchell et al*., 2002]. Spectra of nonalgal particulate absorption, $a_d(\lambda)$, were measured within the same experimental setup as for $a_p(\lambda)$ after

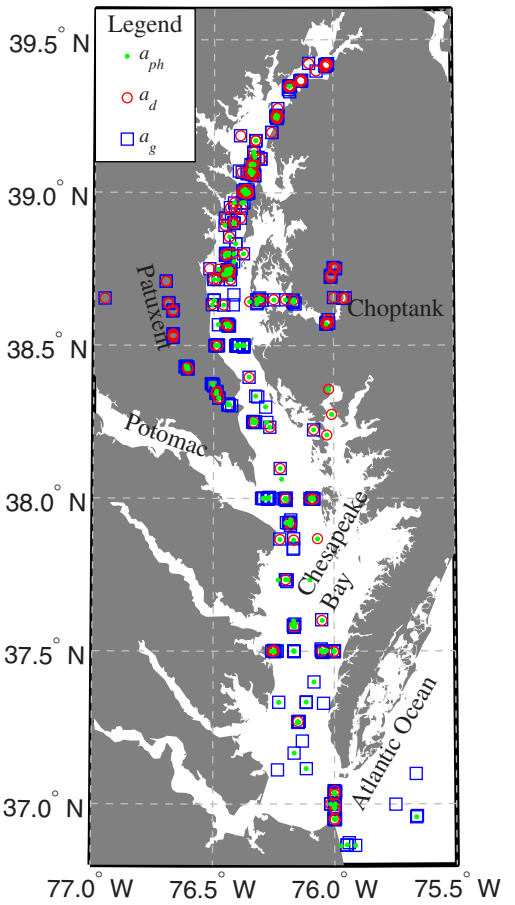

Figure 1. Sampling locations for collecting field data of $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$ in the Chesapeake Bay. The areas in white represent water bodies. Land is shown in gray. Some samples were collected in the tributaries of the bay including the Patuxent, Choptank, and Potomac Rivers.

subjecting the pigments within the particles retained on the filter to hot methanol extraction [*Kishino et al.*, 1985]. The pathlength amplification effect was corrected with an average of two formulas from *Bricaud and Stramski* [1990] and *Nelson and Robertson* [1993] for the GEO-CAPE data and a formula determined for phytoplankton cultures for the remaining data [*Harding et al.*, 2005]. Spectra of CDOM absorption coefficient, $a_g(\lambda)$, were measured in a 10 cm quartz cuvette with a laboratory spectrophotometer using a methodology described by *Mitchell et al.* [2000, 2003]. The absorption coefficients of $a_p(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$ were "null-point" corrected [e.g., *Mitchell et al.*, 2002] so that the magnitude of all spectra in the NIR spectral region (690–710 nm for the GEO-CAPE data set and 720–780 nm for other data sets) is zero. This correction is commonly used because the original measurements with the filter pad method in transmittance configuration are affected by scattering error that can be significant compared with true absorption in the NIR.

After determining the $a_p(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$ coefficients, the total nonwater absorption coefficient, $a_{nw}(\lambda)$, was calculated as $a_p(\lambda) + a_g(\lambda)$. The phytoplankton absorption coefficient, $a_{ph}(\lambda)$, was calculated as $a_p(\lambda) - a_d(\lambda)$. The total nonphytoplankton absorption coefficient, $a_{dg}(\lambda)$, was calculated as $a_d(\lambda) + a_g(\lambda)$.

2.2. Data Quality Control

The absorption data from every station within the original data sets were subject to quality screening to ensure the use of high-quality data in the development of the model and its subsequent evaluation. This screening was conducted for each individual measurement to identify quality issues. All absorption data from the GEO-CAPE data set passed our quality screening. For the LMER-TIES, ACE-INC, and BIOCOMP data sets, a portion of these data was rejected because of issues described below.

For the $a_{ph}(\lambda)$ and $a_d(\lambda)$ data, the rejection was mainly associated with spectral shapes that were deemed to be unreasonable. Specifically, we rejected from the analysis the $a_{ph}(\lambda)$ spectra that exhibited higher magnitude in the red absorption maximum than the blue maximum, because such feature violates typical spectral behavior of phytoplankton pigment absorption [e.g., *Bidigare et al.*, 1990; *Bricaud et al.*, 1999]. In such cases, the data of $a_d(\lambda)$ were also excluded because the erroneous spectral shape of $a_{ph}(\lambda)$ also implies questionable measurement of $a_d(\lambda)$. Another scenario under which the $a_{ph}(\lambda)$ data were rejected is when $a_p(\lambda)$ was so strongly dominated by $a_d(\lambda)$ that the magnitude and spectral shape of $a_{ph}(\lambda)$ was very sensitive to noise in the measured $a_d(\lambda)$. In these cases, however, the $a_d(\lambda)$ data for the same sample could still be considered valid if no other issues were found in the $a_d(\lambda)$ spectrum.

The data quality of $a_d(\lambda)$ spectrum was considered unacceptable if the magnitude of $a_d(\lambda)$ was zero at wavelengths shorter than 650 nm with clear discontinuity between zero and nonzero values. Such a discontinuous feature was possibly caused by unaccounted shift in the baseline associated with the determination of $a_d(\lambda)$ spectrum. Note that for such cases, the $a_{ph}(\lambda)$ data could still be acceptable because the $a_p(\lambda)$ is subject to the same baseline and a possible shift issue was canceled out after the subtraction of $a_d(\lambda)$ from $a_p(\lambda)$.

For the $a_g(\lambda)$ data in the LMER-TIES, ACE-INC, and BIOCOMP data sets, the application of exponential fits to the original data was found necessary to remove small-scale irregularities in the spectral curves associated

with measurement noise. Data from certain spectral regions were excluded from the fitting procedure. For example, data in the spectral range 484–488 nm were always excluded because an apparent instrument artifact (i.e., a positive or negative spike) in the $a_g(\lambda)$ spectrum was frequently observed at or near these wavelengths, likely caused by instrument problems. In addition, large fluctuations in $a_g(\lambda)$ signal were occasionally observed between 570 and 590 nm and the selection or rejection of data within this spectral region was made on a case-by-case basis. Final fitted values of $a_g(\lambda)$ were obtained by fitting the original data at three separate spectral ranges and then matching the three fitted curves together. These ranges were typically 380–470, 430–520, and 490–560 nm. Multiple exponential functions as opposed to one single function were adopted for the fitting because one exponential function was insufficient to fit the original data exhibiting spectral variation in the exponential slope of the measured $a_g(\lambda)$. When noise level in the $a_g(\lambda)$ data was too high to obtain a robust fit, the entire spectrum was rejected.

After the process of quality control, we assembled a final data set consisting of 260 spectra of $a_{ph}(\lambda)$, 207 spectra of $a_d(\lambda)$, and 447 spectra of $a_g(\lambda)$. These data were used to characterize spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$ as well as to determine inequality constraints in the study area as described in detail below. Within this final data set, the $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ coefficients vary by 1 order of magnitude. For example, at $\lambda = 443$ nm, the range of 1st–99th percentiles for a_d, a_g, and a_{ph} is 0.18–3.18, 0.20–3.97, and 0.09–1.28 m^{-1}, respectively. We also note that the final data set contains a subset of 90 stations where valid measurements of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ (i.e., the measurements that passed screening of quality control) are concurrently available. This subset of data was used to evaluate the performance of our model. Within this subset, the contributions of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ to total nonwater absorption $a_{nw}(\lambda)$ are relatively balanced in the short-wavelength portion of the visible spectrum, where magnitudes of these coefficients are generally the highest. For example, the 1st–99th percentiles for the ratios $a_d(443)$:$a_{nw}(443)$, $a_g(443)$:$a_{nw}(443)$, and $a_{ph}(443)$:$a_{nw}(443)$ are 0.23–0.51, 0.18–0.51, and 0.23–0.40, respectively, with no single absorption component exhibiting strong dominance over the other two components.

3. Development of the Stacked-Constraints Model for the Chesapeake Bay

The underlying goal of research presented in this study is to develop a model for partitioning the total nonwater absorption coefficient $a_{nw}(\lambda)$, which satisfies three main requirements. First, the model is to partition $a_{nw}(\lambda)$ into phytoplankton $a_{ph}(\lambda)$, nonalgal particulate $a_d(\lambda)$, and CDOM $a_g(\lambda)$ components, thus including an important capability to discriminate between $a_d(\lambda)$ and $a_g(\lambda)$ which is challenging as the spectral shapes of these coefficients are usually similar. Second, the partitioning problem can be solved with the sole input of $a_{nw}(\lambda)$ so that the model can be applied to data obtained from a wide range of in situ and remote-sensing platforms. Finally, the model shall impose no restrictive assumptions upon the spectral shapes of model-derived $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ to ensure that these optical properties are allowed to vary to the same extent as observed in natural aquatic environments.

This model development effort is essentially a proof-of-concept case study focused on a well-studied and important coastal environment, the Chesapeake Bay. The data collected during historical campaigns (see section 2) are used for model development and evaluation. One essential component of the analysis of field data conducted in this study was a creation of library of representative spectral shapes of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{dg}(\lambda)$ for the Chesapeake Bay (Figure 2). This library was formulated to adequately represent all cases encountered in the study area using a finite manageable number of spectral shapes. It serves as a prerequisite to relax the commonly used assumption in the past studies about the exponential spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$ coefficients. Another important component of the analysis of field data aimed at determining inequality constraints. These constraints define the range of variations associated with spectral shapes of absorption coefficients (Table 1) and are considered only weakly restrictive because the upper and lower boundaries of each constraint were selected typically as extreme values within the range of a given constraint parameter as observed in the study area. The general concept of the determinations and use of inequality constraints in the absorption partitioning models was described in *Zheng and Stramski* [2013a, 2013b] and we follow this concept to develop a new partitioning model in this study.

The overall framework of a new model is similar to the original stacked-constraints model (SCM) [*Zheng and Stramski*, 2013a] although there are a few substantial differences. The original SCM partitions $a_{nw}(\lambda)$ into $a_{dg}(\lambda)$ and $a_{ph}(\lambda)$ with the assumption that the spectral shape of $a_{dg}(\lambda)$ can be described with an

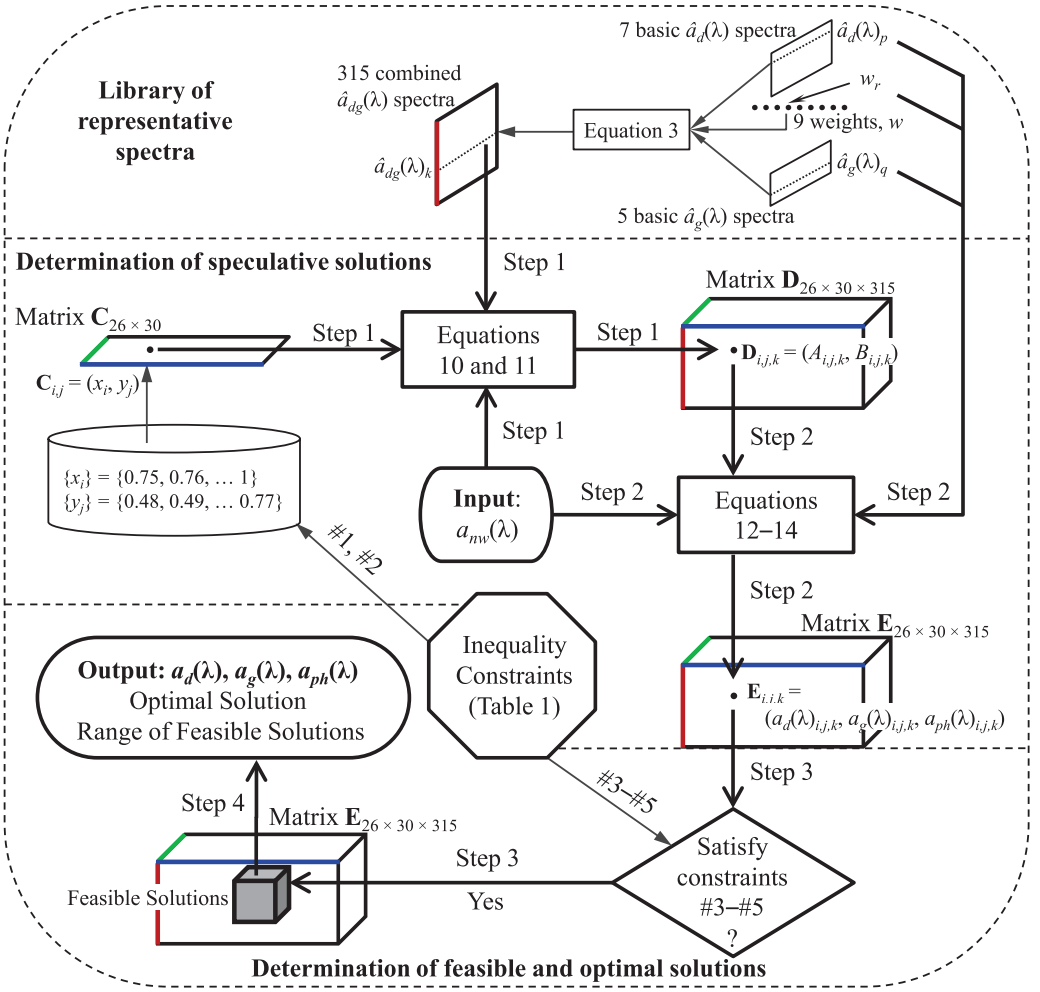

Figure 2. Flowchart of the Generalized Stacked-Constraints Model (GSCM) for partitioning the total light absorption coefficient of natural water (with pure-water contribution subtracted), $a_{nw}(\lambda)$, into phytoplankton, $a_{ph}(\lambda)$, nonalgal particulate, $a_d(\lambda)$, and CDOM, $a_g(\lambda)$, components. Steps 1 and 2 derive speculative solutions, Step 3 identifies feasible solutions among the speculative solutions, and Step 4 determines final model outputs including the optimal solutions for the component absorption coefficients and 10th–90th percentile ranges of feasible solutions for these coefficients. The green and blue lines correspond to the matrix dimensions defined by the number of phytoplankton absorption band ratios, *x* and *y*, respectively. The red lines correspond to the matrix dimension defined by the number of $\hat{a}_{dg}(\lambda)$ spectra. See section 3 for detailed description of the model.

exponential function of wavelength. This assumption, albeit weakly restrictive for the open ocean, can fail in optically more complex waters such as coastal regions rich in absorbing mineral particles or CDOM originating from various sources. In such waters, the spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$ can often depart significantly from a single exponential function which is normally applied over a broad spectral region from near ultraviolet to near infrared [e.g., *Tassan and Ferrari*, 2003; *Babin and Stramski*, 2004; *Stramski et al.*, 2007; *Loiselle et al.*, 2009]. In the present study, the library of spectral shapes derived from field data is incorporated into the stacked-constraints approach, which enables the relaxation of the exponential assumption for $a_{dg}(\lambda)$ as well as the discrimination between $a_d(\lambda)$ and $a_g(\lambda)$. The use of the library generalized the stacked-constraints approach without the need to assume any functional shape for $a_{dg}(\lambda)$. This is yet another major difference between the present model and the original SCM, in which the set of equations referred to as a modified *Bricaud and Stramski* [1990] algorithm is used to solve for the magnitude and slope of the $a_{dg}(\lambda)$ spectrum. In the present model, a slope parameter for $a_{dg}(\lambda)$ is no longer involved but instead the $a_{dg}(\lambda)$ spectrum is described via the library of spectral shapes.

Hereafter we refer to the model developed in this study to as the general stacked-constraints model, GSCM. The flowchart diagram of GSCM is depicted in Figure 2. The GSCM requires input data of total nonwater absorption coefficient, $a_{nw}(\lambda)$, at a minimum of four wavelengths around 412, 443, 490, and 555 nm. These

Table 1. Inequality Constraints Used in the Generalized Stacked-Constraints Model (GSCM) for Partitioning the Total Light Absorption Coefficient of Natural Water (With Pure-Water Contribution Subtracted), $a_{nw}(\lambda)$, into Phytoplankton, $a_{ph}(\lambda)$, Nonalgal Particulate, $a_d(\lambda)$, and CDOM, $a_g(\lambda)$, Components

	Constraints	Physical Meaning
#1	$0.75 < a_{ph}(412)/a_{ph}(443) < 1$	The band ratio of phytoplankton absorption characterizing changes within the short-wavelength portion of the major absorption maximum in the blue spectral region
#2	$0.48 < a_{ph}(490)/a_{ph}(443) < 0.77$	Same as above but within the long-wavelength portion of the blue absorption maximum
#3	$0.76 < a_{ph}(469)/a_{ph}(412) < 1.13$	The band ratio of phytoplankton absorption involving both sides of the blue absorption maximum
#4	$0.19 < a_{ph}(555)/a_{ph}(490) < 0.50$	The band ratio of phytoplankton absorption between the green and cyan spectral regions
#5	$0 < a_d(750)/a_d(443) < 0.3$	The band ratio of nonalgal particulate absorption between the NIR and the blue spectral regions

wavelengths are selected because they coincide closely with the spectral bands available on the past and current satellite ocean color sensors. The spectral resolution of output $a_{ph}(\lambda)$ is the same as that for the model input of $a_{nw}(\lambda)$. The spectral resolution of model-derived $a_d(\lambda)$ and $a_g(\lambda)$ can be as high as 1 nm within the spectral range of 400–750 nm because a number of representative spectral shapes of these coefficients is predefined using hyperspectral field data within that spectral range. Details of the various components of GSCM are described in subsequent sections.

3.1. Library of Spectral Shapes of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{dg}(\lambda)$

The first important task in the development of GSCM is the analysis of large set of field data of $a_d(\lambda)$ and $a_g(\lambda)$ with a purpose of defining a relatively small manageable number of absorption shapes which can be subsequently used to adequately characterize all types of nonalgal particulate and CDOM spectra encountered in the study area. To accomplish this task, we used the so-called hierarchical cluster analysis (HCA) [e.g., *Anderberg*, 1973], which essentially compares the similarity in the shape among all spectra in the field data and assembles the sets of similar spectra within separate clusters. We applied this approach to the field data obtained in the Chesapeake Bay comprising 207 spectra of $a_d(\lambda)$ and 447 spectra of $a_g(\lambda)$ (see section 2). Only data within the spectral range of 400–750 nm were included in this analysis. The shorter wavelengths were excluded because the pathlength amplification effect involved in the $a_p(\lambda)$ and $a_d(\lambda)$ measurements is not well characterized in the UV.

To facilitate a comparative analysis of spectral shapes, we normalized the spectral values within each absorption spectrum of $a_d(\lambda)$ and $a_g(\lambda)$ by the sum of all spectral values between 400 and 750 nm so that the integral of each normalized spectrum, hereafter denoted by $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$, is equal to unity:

$$\hat{a}_d(\lambda) = a_d(\lambda) \Big/ \sum_{\lambda=400}^{750} a_d(\lambda) \tag{1}$$

$$\hat{a}_g(\lambda) = a_g(\lambda) \Big/ \sum_{\lambda=400}^{750} a_g(\lambda) \tag{2}$$

The normalized spectra, $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$, were categorized into 7 and 5 clusters, respectively, using the HCA. The number of clusters was determined to achieve the lowest similarity between neighboring clusters relative to the similarity among individual spectra within each cluster. The similarity is defined to be inversely proportional to the Euclidean distance between the vectors of normalized absorption spectra, where the vector is a set of all spectral values within the normalized spectrum. The final spectra of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ representative of each cluster were calculated by averaging all spectra within each cluster. As a result a total of seven $\hat{a}_d(\lambda)$ spectra (Figure 3a) and five $\hat{a}_g(\lambda)$ spectra (Figure 3b) were defined. The tabulated values of these spectra are provided in supporting information (Table S1). We consider these basic representative spectra of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ as proxies that adequately capture the complete variability in the spectral shapes of nonalgal particulate $a_d(\lambda)$ and CDOM $a_g(\lambda)$ absorption coefficients in the Chesapeake Bay.

Assuming that these basic spectral shapes are generally representative of the true absorption properties of the Chesapeake Bay waters, we can address the partitioning problem by solving for the magnitudes of

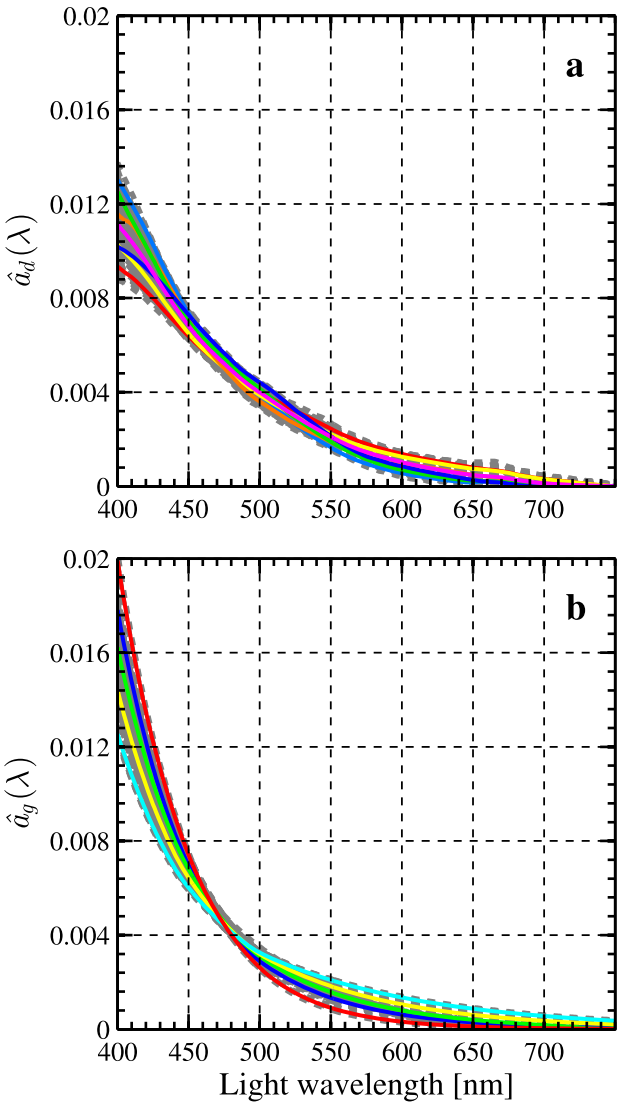

Figure 3. (a) Seven representative normalized absorption spectra of nonalgal particles $\hat{a}_d(\lambda)$ and (b) five representative normalized absorption spectra of colored dissolved organic matter $\hat{a}_g(\lambda)$, which characterize the variability in the spectral shape of these absorption components in the Chesapeake Bay. These representative spectra are plotted with different colors and each spectrum was normalized by the integral over the 400–750 nm spectral range. The gray dashed curves shown as a background for the colored curves were obtained from the original measurements of nonalgal particulate (207 spectra) and CDOM absorption (447 spectra) coefficients.

$a_d(\lambda)$ and $a_g(\lambda)$ for each and every combination of the basic spectral shapes $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$, which requires the consideration of a total of 35 (7×5) scenarios. However, there is a complication which is associated with experimental limitations in the measurements of $a_d(\lambda)$. Specifically, all field data of $a_d(\lambda)$ used in this study were processed with the "null-point" correction because the measured absorption signal in the NIR obtained with the filter-pad T-method was potentially subject to scattering errors and/or other artifacts of the method. As a result, the 7 basic spectra of $\hat{a}_d(\lambda)$ derived from the field data do not resolve the potential nonzero magnitude and variations in the NIR absorption by nonalgal particles. The lack of such a capability leads to a problem for the Chesapeake Bay because this water body receives significant terrigenous input of mineral particles, which can likely absorb in the NIR [*Tassan and Ferrari*, 2003; *Babin and Stramski*, 2004; *Stramski et al*., 2004, 2007; *Tzortziou et al*., 2006; *Röttgers et al*., 2014]. Therefore, it is critical for our model to have the ability to account for potential NIR absorption of nonalgal particles, a feature that is missing in the basic spectral shapes of $\hat{a}_d(\lambda)$.

To account for unknown magnitude of $a_d(\lambda)$ in the NIR in the present model, we introduce a variable, B, which is a spectrally independent offset of $a_d(\lambda)$. This offset is meant to represent true absorption of nonalgal particulate matter in the NIR. Model-derived $a_d(\lambda)$ is calculated by adding the offset B to all spectral values of nonalgal particulate absorption derived from the basic $\hat{a}_d(\lambda)$ spectra. Such procedure annuls the "null-point" correction in which a constant value was subtracted from the entire spectrum of $a_d(\lambda)$ data that were obtained from measurements with the filter-pad T-method. The introduction of variable B is accompanied by undesirable increase in the number of unknowns for solving the partitioning problem. To reduce the total number of unknowns, we create an intermediate step in model calculations in which the two separate magnitudes of $a_d(\lambda)$ and $a_g(\lambda)$ are represented in terms of one scaling parameter, A, which denotes the magnitude of $a_{dg}(\lambda)$. The proportions of $a_d(\lambda)$ and $a_g(\lambda)$ within $a_{dg}(\lambda)$ are represented using an array of nine discrete weighting factors, i.e., $w = \{0.1, 0.2, \ldots, 0.9\}$. For any given measurement of water sample, the values of the weighting factors specific to $a_d(\lambda)$ and $a_g(\lambda)$ are dependent upon each other so that their sum always equals to 1. The values of 0 and 1 are excluded from candidate weighting factors which results from the assumption that nonalgal particles and CDOM are always present in water. We also note that the use of

higher resolution than 0.1 for the discrete weighting factors is in principle possible but was deemed unessential in this study from the standpoint of model performance.

With the addition of the weighting factors, the total number of scenarios for the parameterization of $a_{dg}(\lambda)$ increases from 35 to 315 ($7 \times 5 \times 9$) to represent all combinations of various basic spectral shapes of $\hat{a}_d(\lambda)$, $\hat{a}_g(\lambda)$, and the discrete values of weighting factors. For each scenario, the spectral shape of $a_{dg}(\lambda)$ before the NIR absorption offset B is added, denoted as $\hat{a}_{dg}(\lambda)$, can be calculated as a linear combination of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ according to (see also Figure 2):

$$\hat{a}_{dg}(\lambda)_k = w_r \hat{a}_d(\lambda)_p + (1 - w_r)\hat{a}_g(\lambda)_q \tag{3}$$

where

$\hat{a}_{dg}(\lambda)_k$ kth spectrum of $\hat{a}_{dg}(\lambda)$, $k = 1, 2, \ldots, 315$;

$\hat{a}_d(\lambda)_p$ pth basic spectrum of $\hat{a}_d(\lambda)$, $p = 1, 2, \ldots, 7$;

$\hat{a}_g(\lambda)_q$ qth basic spectrum of $\hat{a}_g(\lambda)$, $q = 1, 2, \ldots, 5$;

w_r rth element of the array of weighting factors w, $r = 1, 2, \ldots, 9$.

Each of 315 spectra of $\hat{a}_{dg}(\lambda)$ corresponds to a unique combination of one basic spectrum of $\hat{a}_d(\lambda)$, one basic spectrum of $\hat{a}_g(\lambda)$, and one value of weighting factor w. The 315 spectra of $\hat{a}_{dg}(\lambda)$, 7 spectra of $\hat{a}_d(\lambda)$, and 5 spectra of $\hat{a}_g(\lambda)$ constitute a library of spectral shapes for total nonphytoplankton absorption, nonalgal particulate absorption, and CDOM absorption, and they are all used in the calculations of the model (Figure 2).

3.2. Inequality Constraints

In addition to determining the spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$, the field data from the Chesapeake Bay are used also to establish a set of inequality constraints which are a key part of our modeling formalism (Table 1). The constraint variables are defined in terms of band ratios of phytoplankton absorption $a_{ph}(\lambda)$ and nonalgal particulate absorption $a_d(\lambda)$ with a purpose to capture specific spectral features of these absorption coefficients. The constraints #1 and #2 are defined as band ratios of $a_{ph}(\lambda)$ describing changes in phytoplankton absorption on both sides of the major absorption maximum in the blue spectral region. The constraint #3 ensures that the derived $a_{ph}(\lambda)$ assumes reasonable values at both sides of the maximum in the blue spectral region. The wavelengths of 412 and 469 nm involved in this constraint are located more or less symmetrically relative to the absorption maximum. We note that in the case when the input data of $a_{nw}(\lambda)$ are not available at 469 nm, the value of model-derived $a_{ph}(469)$ can be estimated from interpolation between the wavelengths of 443 and 490 nm, which are always required to run the model. The constraint #4 ensures realistic values for the ratio of derived $a_{ph}(\lambda)$ at 490 and 555 nm, which implies a reasonable spectral behavior of $a_{ph}(\lambda)$ within the long-wavelength tail of blue absorption maximum. Variability in the values of constraint variables #1–#4 in the Chesapeake Bay is demonstrated by histograms based on a comprehensive set of 260 field measurements of $a_{ph}(\lambda)$ (Figure 4). The upper and lower bounds for inequality constraints #1–#4 were determined as the 1st and 99th percentiles of the histograms, which is consistent with the approach used in the original SCM [*Zheng and Stramski*, 2013a].

The constraint #5 ensures that the nonalgal particulate absorption spectrum has a reasonable magnitude in the NIR relative to the blue spectral region. This constraint is introduced to account for possible significant absorption by particles in the NIR, which is expected mainly in waters rich in mineral particles [e.g., *Tassan and Ferrari*, 2003; *Babin and Stramski*, 2004; *Bowers and Binding*, 2006; *Stramski et al*., 2007; *Röttgers et al*., 2014]. The upper boundary of constraint #5 is determined from this literature because the available field data from the Chesapeake Bay provide no useful information about the NIR particulate absorption. *Tassan and Ferrari* [2003] showed that the ratio of absorbance between 750 and 440 nm may range between 0.04 and 0.31 for marine and freshwater environments. *Stramski et al*. [2004] showed that the NIR (750 nm) to blue (400 nm) absorption ratio of Asian dust particles ranges from very low values close to 0 to about 0.1. Similar ratio between 800 and 400 nm for various mineral-dominated particle assemblages was reported in the range of 0.032–0.15 [*Babin and Stramski*, 2004] and 0.025–0.23 [*Stramski et al*., 2007]. Measurements on water samples collected in several marine coastal and estuarine regions suggest that the upper limit of the NIR (750 nm) to blue (400 nm) absorption ratio is about 0.3 [*Röttgers et al*., 2014]. On the basis of these data, a value of 0.3 is here considered as an appropriate upper limit of constraint #5. We note that the derivation

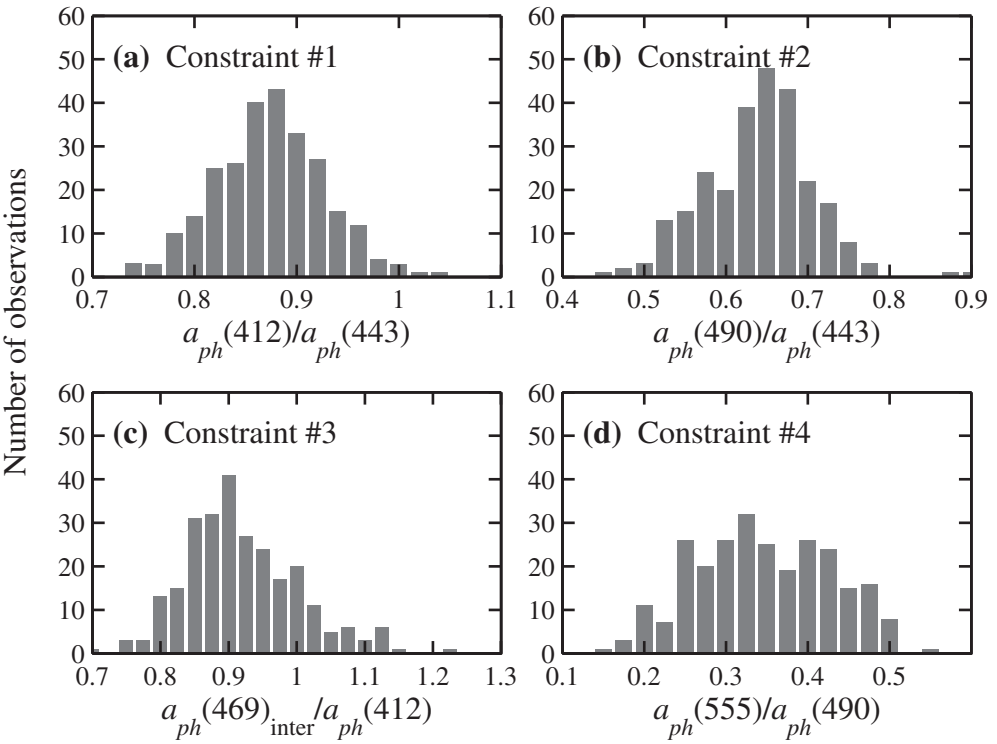

Figure 4. Histograms of parameters involved in the inequality constraints #1–#4 as obtained from 260 field measurements of phytoplankton absorption coefficient, $a_{ph}(\lambda)$, taken in the Chesapeake Bay. The $a_{ph}(469)_{inter}$ denotes the linearly interpolated value between $a_{ph}(443)$ and $a_{ph}(490)$.

of $a_d(\lambda)$ at 750 nm which is required for constraint #5 does not necessarily require the availability of input data of $a_{nw}(\lambda)$ at 750 nm. This derivation can be accomplished with input of $a_{nw}(\lambda)$ provided for a minimum set of four wavelengths required by our model, i.e., 412, 443, 490, and 555 nm. This is because $a_d(\lambda)$ is derived with high spectral resolution as defined by the representative spectra of $\hat{a}_d(\lambda)$ which were obtained from hyperspectral absorption measurements.

3.3. Solving for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$

Having determined a set of representative spectra of $\hat{a}_d(\lambda)$, $\hat{a}_g(\lambda)$, and $\hat{a}_{dg}(\lambda)$, as well as a set of inequality constraints, our model is equipped with information necessary to solve for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$. The solutions of the model can be obtained with input data of $a_{nw}(\lambda)$ provided for a minimum set of four basic wavelengths in the blue and green spectral regions, i.e., 412, 443, 490, and 555 nm. In this particular case of spectrally limited input data our model finds solutions of $a_{ph}(\lambda)$ at these four input wavelengths, but the solutions of $a_d(\lambda)$ and $a_g(\lambda)$ are obtained at high spectral resolution consistent with the resolution of predefined representative spectra of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$. If input data of $a_{nw}(\lambda)$ are available at higher spectral resolution, the solutions of $a_{ph}(\lambda)$ are also found with correspondingly high spectral resolution. The model operates via several steps starting with calculations of many speculative solutions, followed by the determinations of a smaller set of feasible solutions, and finally the identification of optimal solutions. These steps are described below.

3.3.1. Determination of Speculative Solutions

The purpose of the first two steps of GSCM is to calculate a large number of speculative solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ using the input of spectral values of $a_{nw}(\lambda)$. Major difference in these calculations between the GSCM and the original SCM is associated with parameterization of the spectral shape of nonphytoplankton absorption coefficient, $a_{dg}(\lambda)$. Whereas the original SCM adopts an exponential function of wavelength, the GSCM uses 315 realistic shapes of $\hat{a}_{dg}(\lambda)$ which include the possibility of nonzero magnitude of nonalgal particulate absorption in the NIR as accounted for by the parameter B. As emphasized above, another major advancement of GSCM is the separation of nonalgal particulate absorption $a_d(\lambda)$ and CDOM absorption $a_g(\lambda)$, which is not accomplished by the original SCM.

Step 1 (Figure 2) derives speculative solutions for two unknowns, i.e., the magnitude or scaling parameter A of the $a_{dg}(\lambda)$ spectrum and the NIR offset parameter B of the $a_d(\lambda)$ spectrum, using the library of 315 shapes of $\hat{a}_{dg}(\lambda)$ and constraints #1 and #2 (Table 1). The calculations start with two phytoplankton absorption band ratios, $x \equiv a_{ph}(412):a_{ph}(443)$ and $y \equiv a_{ph}(490):a_{ph}(443)$. The possible values of the two ratios are represented by two sets of evenly distributed numbers with boundaries defined by constraints #1 and #2, i.e., $x = \{0.75, 0.76, \ldots, 1\}$ and $y = \{0.48, 0.49, \ldots, 0.77\}$, respectively (Figure 2). All possible combinations for the pair of x_i and y_j can be represented by a 26×30 matrix **C**, where the subscripts i and j identify one possible pair of x and y values (Figure 2). For a given pair of x_i and y_j, a given shape of $\hat{a}_{dg}(\lambda)_k$, and given spectral values of $a_{nw}(\lambda)$ at 412, 443, and 490 nm, the system of equations can be written as:

$$x_i = a_{ph}(412)_{i,j,k} / a_{ph}(443)_{i,j,k} \tag{4}$$

$$y_j = a_{ph}(490)_{i,j,k} / a_{ph}(443)_{i,j,k} \tag{5}$$

$$a_{nw}(\lambda) = a_{ph}(\lambda)_{i,j,k} + a_{dg}(\lambda)_{i,j,k} \tag{6}$$

$$a_{dg}(\lambda)_{i,j,k} = A_{i,j,k}\hat{a}_{dg}(\lambda)_k + B_{i,j,k} \tag{7}$$

Combining equations (4–7), we obtain a system of two equations with two unknowns, i.e., the scaling factor $A_{i,j,k}$ of the $\hat{a}_{dg}(\lambda)$ spectrum and the NIR parameter $B_{i,j,k}$ of the $a_d(\lambda)$ spectrum:

$$a_{nw}(412) - x_i a_{nw}(443) = A_{i,j,k}[\hat{a}_{dg}(412)_k - x_i \hat{a}_{dg}(443)_k] + B_{i,j,k}(1 - x_i) \tag{8}$$

$$a_{nw}(490) - y_j a_{nw}(443) = A_{i,j,k}[\hat{a}_{dg}(490)_k - y_j \hat{a}_{dg}(443)_k] + B_{i,j,k}(1 - y_j) \tag{9}$$

The speculative solutions for the two unknowns, $A_{i,j,k}$ and $B_{i,j,k}$, are then derived as:

$$A_{i,j,k} = \frac{(1-x_i)[a_{nw}(490) - y_j a_{nw}(443)] - (1-y_j)[a_{nw}(412) - x_i a_{nw}(443)]}{(1-x_i)[\hat{a}_{dg}(490)_k - y_j \hat{a}_{dg}(443)_k] - (1-y_j)[\hat{a}_{dg}(412)_k - x_i \hat{a}_{dg}(443)_k]} \tag{10}$$

$$B_{i,j,k} = \frac{a_{nw}(412) - x_i a_{nw}(443) - A_{i,j,k}[\hat{a}_{dg}(412)_k - x_i \hat{a}_{dg}(443)_k]}{1 - x_i} \tag{11}$$

These calculations are repeated with the same input data of $a_{nw}(\lambda)$ for each and every combination of x_i, y_j, and $\hat{a}_{dg}(\lambda)_k$. As a result, a complete set of speculative solutions for A and B is obtained, which is represented by matrix **D** with dimensions of $26 \times 30 \times 315$ (see Figure 2).

Step 2 derives speculative solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ using the basic shapes of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ as well as the speculative solutions of A and B derived in Step 1. For each speculative solution of $A_{i,j,k}$ and $B_{i,j,k}$, the three component absorption coefficients are calculated as:

$$a_d(\lambda)_{i,j,k} = A_{i,j,k} w_r \hat{a}_d(\lambda)_p + B_{i,j,k} \tag{12}$$

$$a_g(\lambda)_{i,j,k} = A_{i,j,k}(1 - w_r)\hat{a}_g(\lambda)_q \tag{13}$$

$$a_{ph}(\lambda)_{i,j,k} = a_{nw}(\lambda) - a_d(\lambda)_{i,j,k} - a_g(\lambda)_{i,j,k} \tag{14}$$

Note that the indices p, q, and r can be traced back to those used in the calculations of $\hat{a}_{dg}(\lambda)_k$ spectra from equation (3). The calculations with equations (12–14) are repeated with the same input data of $a_{nw}(\lambda)$ for each speculative solution of $A_{i,j,k}$ and $B_{i,j,k}$. As a result, a complete set of speculative solutions for $a_d(\lambda)_{i,j,k}$, $a_g(\lambda)_{i,j,k}$, and $a_{ph}(\lambda)_{i,j,k}$ is obtained, which is represented by matrix **E** with dimensions of $26 \times 30 \times 315$ (Figure 2).

3.3.2. Determination of Feasible and Optimal Solutions

The next two steps (3 and 4) of the model determine feasible solutions from the large number of speculative solutions, and identify final optimal solutions within the range of feasible solutions. The feasible solutions are also used to characterize the range of variation for most likely solutions around the selected optimal solution. As these two steps of GSCM are conceptually the same as those in the original SCM, we here provide only a brief description. More detailed description is in *Zheng and Stramski* [2013a].

Step 3 of GSCM reduces the large set of speculative solutions to a smaller subset of feasible solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$. This domain of feasible solutions is determined by applying additional three

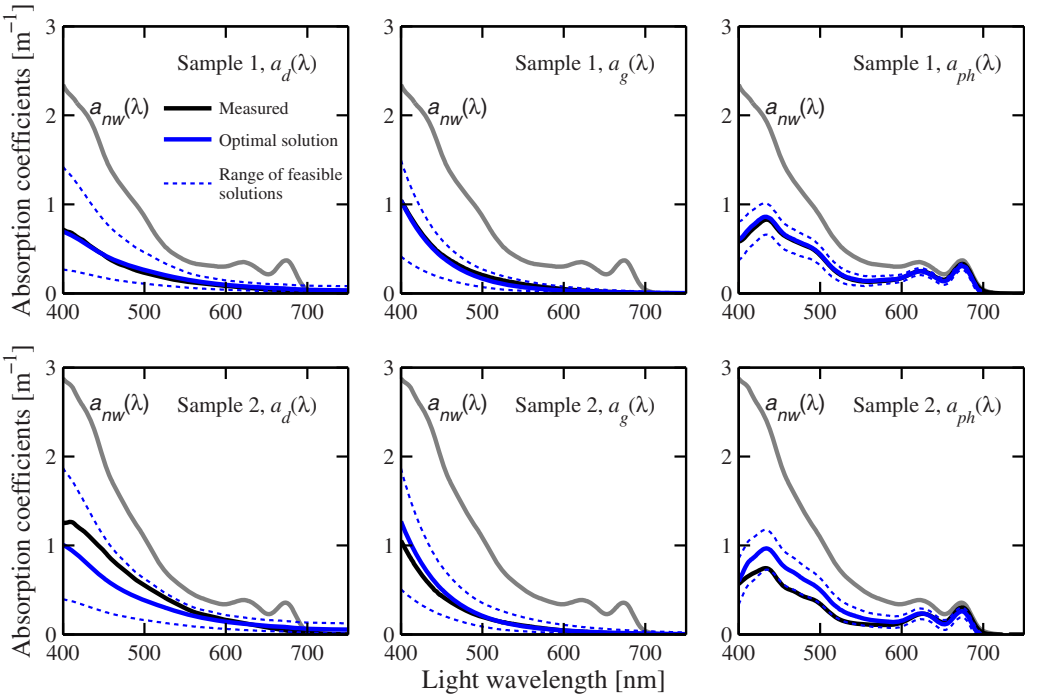

Figure 5. Example comparisons of measured absorption coefficients and the results obtained with the partitioning model for surface water samples collected in the Chesapeake Bay. An example of the (top) "best-case-scenario" and (bottom) "worst-case-scenario" in terms of agreement between model-derived and measured values is given.

inequality constraints (i.e., the constraints #3–#5 in Table 1) to every case of the speculative solutions. The feasible solutions are identified as those that concurrently satisfy all inequality constraints. The last step of the model (Step 4 in Figure 2) generates the final outputs which include the optimal solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ which are selected from feasible solutions and a range characterizing variation of most likely solutions within the domain of all feasible solutions. The optimal solutions for each data product are determined in terms of the median values of feasible solutions at each wavelength separately. The use of median values as optimal solutions was shown to provide good overall performance of the original SCM [*Zheng and Stramski*, 2013a]. The range of variation of feasible solutions around the optimal solutions is characterized by the 10th and 90th percentiles of all feasible solutions. Such range was shown to ensure sufficiently high probability that the actual measured values fall within this range [*Zheng and Stramski*, 2013a].

4. Evaluation of the Performance of the Model

In this section, we evaluate the performance of the model by comparing the $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ absorption spectra derived from the model with those obtained from the measurements. For this purpose, we ran the GSCM with input data of $a_{nw}(\lambda)$ obtained from 90 field measurements of absorption coefficients $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ in the Chesapeake Bay, as described in section 2. Several features of the data set comprising these 90 measurements support its suitability for evaluating the model performance. The boundaries of the inequality constraints can be considered largely independent of the field data because the constraints #1–#4 correspond to only a few extreme cases in the field data set and the boundaries of constraint #5 were determined from literature. Whereas the basic spectra of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ in our model represent the variability in the spectral shape of these coefficients, the weighting factors that determine the proportions of $\hat{a}_d(\lambda)$ and $\hat{a}_g(\lambda)$ in $\hat{a}_{dg}(\lambda)$ are entirely independent of the field data, so the resultant spectral shapes of $\hat{a}_{dg}(\lambda)$ can be also considered largely independent of the field data. Furthermore, the data set comprising 90 measurements used for evaluating the model represent only a small portion of the entire data set used for developing the model, i.e., only stations where all three component absorption coefficients concurrently passed our data quality screening are used for the evaluation of the model. Below we

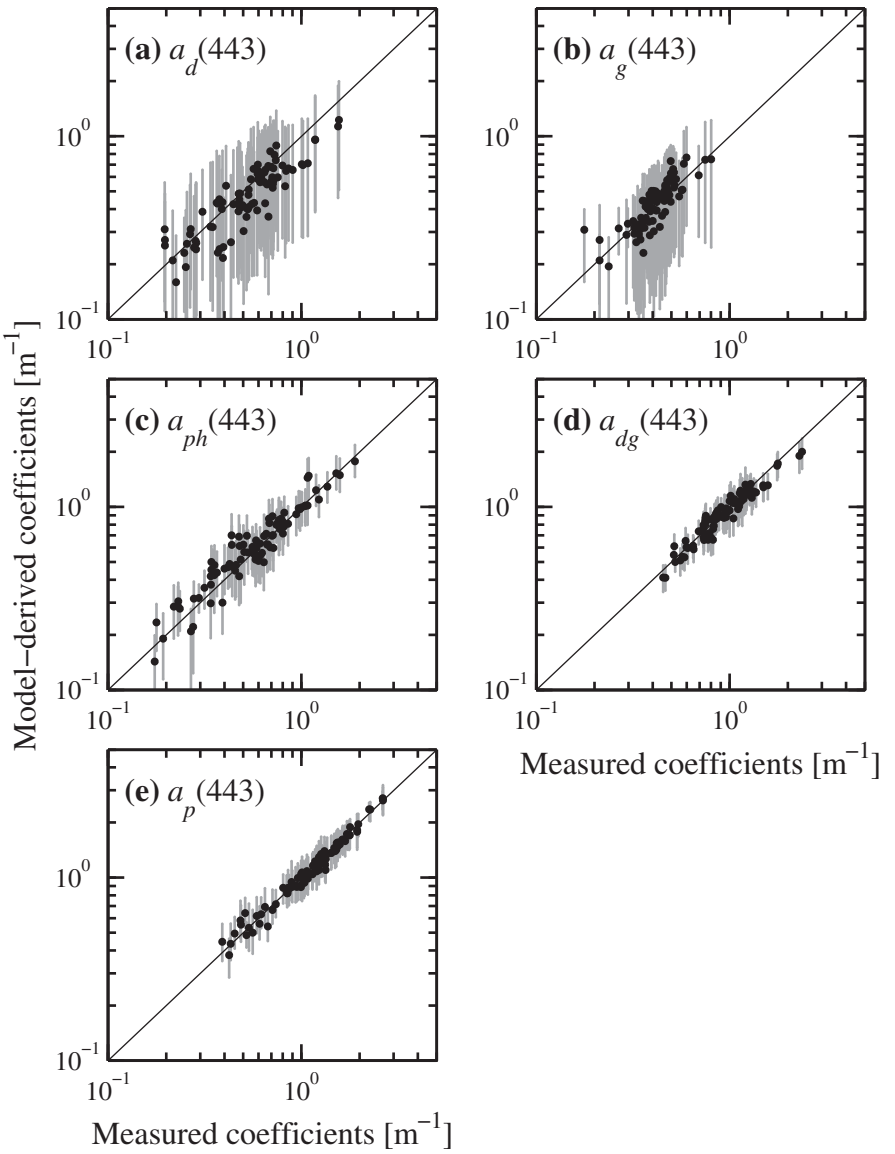

Figure 6. Comparisons of model-derived and measured values of nonalgal particulate absorption coefficient, $a_d(\lambda)$, CDOM absorption coefficient, $a_g(\lambda)$, phytoplankton absorption coefficient, $a_{ph}(\lambda)$, nonphytoplankton absorption coefficient, $a_{dg}(\lambda)$, and total particulate absorption coefficient, $a_p(\lambda)$, at a light wavelength of 443 nm. Black solid circles indicate the optimal solutions of the model, which were obtained with the use of input data of measured total nonwater absorption coefficients, $a_{nw}(\lambda)$, for the data set consisting of 90 samples from the Chesapeake Bay (see text for more details). Gray vertical bars represent the 10th–90th percentile ranges of feasible solutions for all presented cases. The 1:1 relationship is shown as the diagonal line.

evaluate the model with two sets of input data of $a_{nw}(\lambda)$; first, the original data set of 90 measurements, and second, the modified data set that simulates the possible absorption of nonalgal particles in the NIR spectral region.

4.1. Evaluation Based on the Original Field Data of $a_{nw}(\lambda)$

The values of component absorption coefficients, including the nonalgal particulate component $a_d(\lambda)$, in the original data set collected in the Chesapeake Bay is zero or close to zero in the NIR as a result of "null-point" correction applied to the measurements (see section 2). We here describe results of evaluation of our model for a subset of 90 measurements from this original data set. The most important outputs of the model are the optimal (i.e., the median-based) solutions and the 10th–90th percentile range of feasible solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ at all wavelengths of input $a_{nw}(\lambda)$. Example results are shown for two

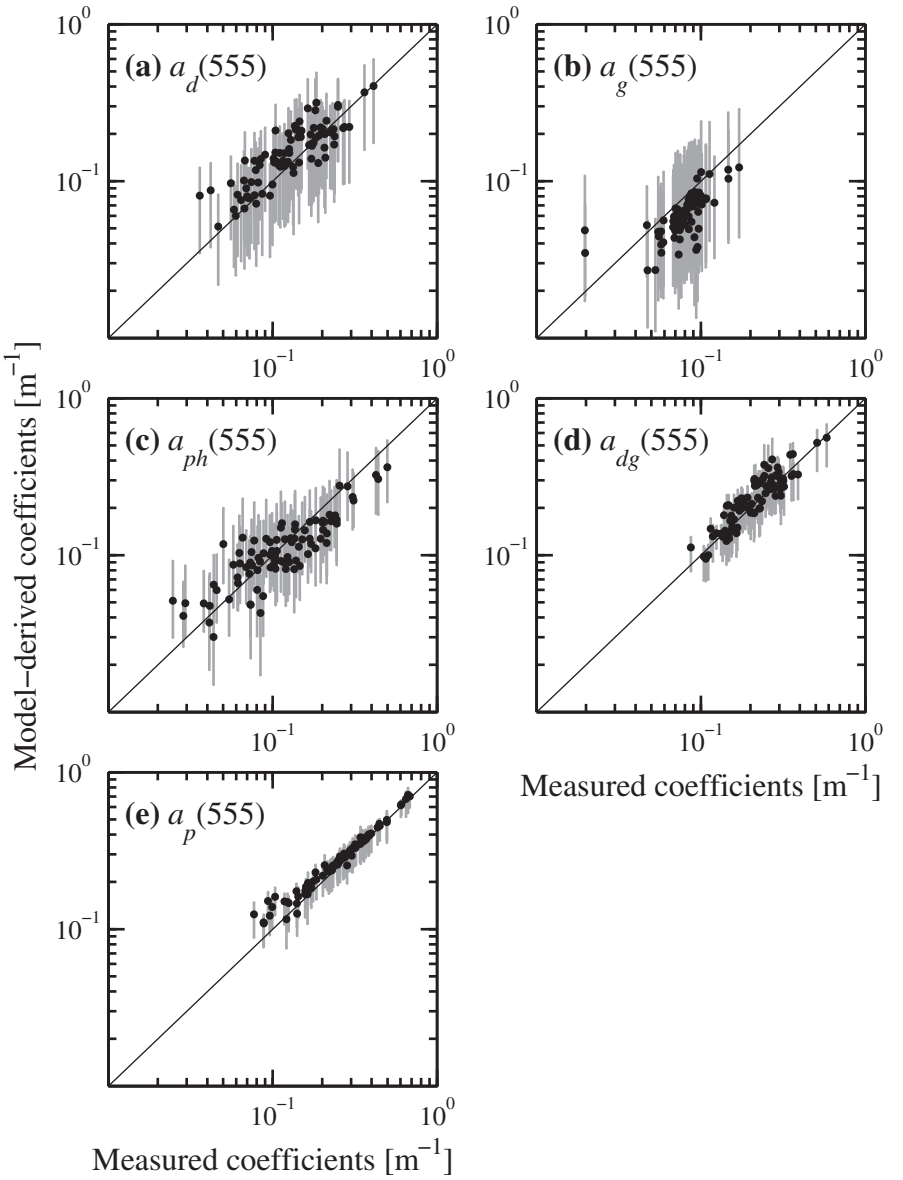

Figure 7. Same as Figure 6 but for the light wavelength of 555 nm.

samples representing the "best-case-scenario" (Figure 5, top) and "worst-case-scenario" (Figure 5, bottom) in terms of the agreement between the model-derived and measured spectra. Although the model can operate with input $a_{nw}(\lambda)$ at only four wavelengths, we show full spectra with high spectral resolution which is consistent with the input data of $a_{nw}(\lambda)$. For both scenarios, the optimal solutions of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ agree well with the measured spectra and the 10th–90th percentile ranges encompass the measured spectra. The optimal solutions match very well the measured spectra for the "best-case-scenario," and still agree reasonably well with measurements for the "worst-case-scenario."

We now turn to comparative analysis of model-derived and measured values of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ for the entire data set of 90 concurrent measurements of these coefficients. The nonphytoplankton absorption coefficient $a_{dg}(\lambda)$, calculated as $a_d(\lambda) + a_g(\lambda)$, and the total particulate absorption coefficient $a_p(\lambda)$, calculated as $a_d(\lambda) + a_{ph}(\lambda)$, are also included in this comparison. Figures 6 and 7 show results from this analysis for the blue (443 nm) and green (555 nm) wavelengths, respectively. The selection of these wavelengths is

Table 2. Summary of Error Statistics for Selected Output Variables of GSCM[a]

Variable	*R*	*MR*	*SIQR* (%)	*MPD* (%)	*RMSD* (m^{-1})
$a_d(412)$	0.906	0.874	13.18	15.31	0.179
$a_g(412)$	0.841	1.080	10.95	13.79	0.133
$a_{ph}(412)$	0.941	1.019	13.32	11.92	0.097
$a_d(443)$	0.896	0.913	12.62	17.15	0.139
$a_g(443)$	0.836	1.064	11.12	12.74	0.072
$a_{ph}(443)$	0.956	1.056	11.77	10.90	0.100
$a_d(490)$	0.869	1.014	17.80	17.74	0.084
$a_g(490)$	0.829	0.960	9.75	8.76	0.032
$a_{ph}(490)$	0.946	1.013	14.11	13.19	0.076
$a_d(555)$	0.833	1.140	19.88	19.27	0.045
$a_g(555)$	0.768	0.787	7.51	22.33	0.023
$a_{ph}(555)$	0.914	0.921	22.16	23.83	0.044

[a]The value of *R* is the correlation coefficient between the model-derived and measured values. *MR* is the median ratio of model-derived to measured values, and *SIQR* is the semiinterquartile range for this ratio expressed in percent and calculated as $SIQR = 100\,(QR_3 - QR_1)/2$, where QR_1 is the first quartile and QR_3 is the third quartile of this ratio. The *MPD* is the median absolute percent difference calculated as the median of the individual absolute percent differences $PD_i = 100\,|Y_i - X_i|/X_i$ where Y_i are the model derived and X_i are the measured values. The *RMSD* is the root mean square deviation between the model-derived and measured values calculated as $RMSD = [N^{-1}\sum(Y_i - X_i)^2]^{1/2}$, with the summation from $i = 1$ to N. N is the number of observations used for deriving the error statistics. $N = 90$. The model-derived values involved in the calculations of error statistics refer to the optimal solutions of our model.

adequate for general characterization of model performance at a few major spectral bands that are available on past and current satellite ocean color sensors. These figures compare the optimal solutions and 10th–90th percentile ranges derived from the model with the measured values for $a_d(\lambda)$, $a_g(\lambda)$, $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $a_p(\lambda)$.

Assuming that the differences between the model-derived and measured absorption coefficients can be considered to represent errors of the model, we calculated several error parameters for evaluating model performance. Specifically, the median of the ratio of model-derived optimal solutions to measured values, *MR*, was calculated to provide a measure of overall bias in the modeled data relative to measurements. The semiinterquartile range, *SIQR*, calculated for the ratio of modeled derived to measured values indicates the spread of the modeled data. The random errors are quantified in terms of the median value of the absolute percent difference, *MPD*, between the model-derived and measured data, and also by the root mean square deviation, *RMSD*, between these data. The correlation coefficient, *R*, provides additional information on how well the model-derived data agree with the measurements over their dynamic range. All these error statistics are listed in Table 2 for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ at four basic wavelengths required as input of the model, namely 412, 443, 490, and 555 nm. The formulas for calculating the error parameters are also given in Table 2.

Figures 6 and 7 as well as Table 2 show that our model performs generally well in terms of partitioning the total nonwater absorption $a_{nw}(\lambda)$ into nonalgal particulate $a_d(\lambda)$, CDOM $a_g(\lambda)$, and phytoplankton $a_{ph}(\lambda)$ components. At the blue wavelength of 443 nm, the optimal solutions for $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ are generally in good agreement with measured values (Figures 6a–6c). The systematic component of the error in model-derived $a_d(443)$, $a_g(443)$, and $a_{ph}(443)$ represented by a departure of *MR* from 1 are within $\pm$10% (Table 2). The random components of the error represented by *MPD* or *RMSD* are also generally small. The *MPD* values for these three absorption coefficients range between 11% and 17% and *RMSD* between 0.07 and 0.14 m^{-1} (Table 2). Compared with the performance of the model for $a_d(443)$, $a_g(443)$, and $a_{ph}(443)$, better agreement is found between the model-derived and measured values of $a_{dg}(443)$ and $a_p(443)$ (Figures 6d and 6e). This can be attributed to the enhanced absorption signal relative to noise level for these two coefficients. This analysis for the blue band where phytoplankton exhibits maximum absorption indicates that the various absorption components are derived with reasonably good accuracy by our model.

At the green wavelength of 555 nm, larger errors are observed compared with the blue band. For example, the systematic errors calculated from *MR* are +14%, −21%, and −8% for $a_d(555)$, $a_g(555)$, and $a_{ph}(555)$, respectively (Table 2). The random errors for these coefficients based on *MPD* are 19%, 22%, and 24%, respectively (Table 2). The larger errors in the green portion of the spectrum are associated with generally low or minimum magnitudes of these absorption coefficients at this wavelength.

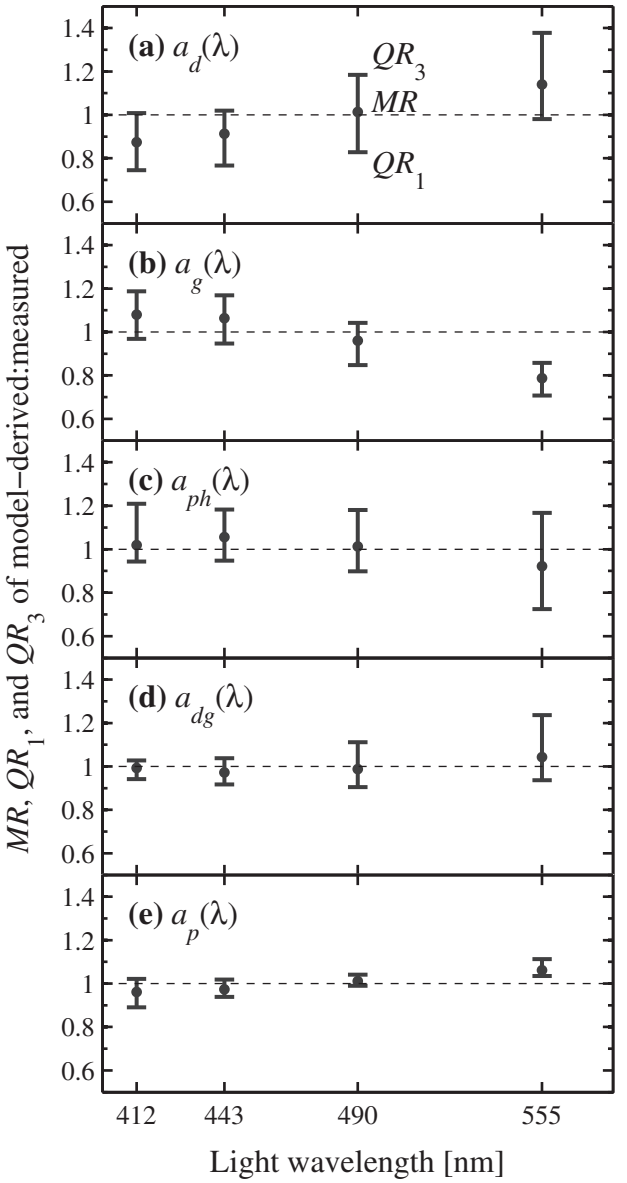

Figure 8. (a) The spectral dependence of the median ratio, *MR* (solid circles), of the model-derived optimal solution to the measured value of the nonalgal particulate absorption coefficient, $a_d(\lambda)$, as determined by the application of the model to the data set of 90 absorption measurements from the Chesapeake Bay. The upper and lower bounds of the error bar represent the first and third quartile ratios, QR_1 and QR_3, of the model-derived to measured values of $a_d(\lambda)$. The plots below are the same as Figure 8a but for (b) the CDOM absorption coefficient, $a_g(\lambda)$, (c) the phytoplankton absorption coefficient, $a_{ph}(\lambda)$, (d) the nonphytoplankton absorption coefficient, $a_{dg}(\lambda)$, and (e) the total particulate absorption coefficient, $a_p(\lambda)$. The thin horizontal dashed lines corresponding to the ordinate of 1 indicate no bias of the model-derived values relative to the measured values.

In addition to the optimal solutions, Figures 6 and 7 show the range of feasible solutions as defined by the 10th and 90th percentiles. The vertical bars depicting the range of feasible solutions cross the 1:1 lines in most cases, indicating that this range effectively characterizes the uncertainties of model-derived solutions. The length of the vertical bars representing the 10th and 90th inter-percentile range is generally larger for $a_d(\lambda)$ and $a_g(\lambda)$ than for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $a_p(\lambda)$. This indicates that the partitioning results for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $a_p(\lambda)$ are generally subject to smaller uncertainties than $a_d(\lambda)$ and $a_g(\lambda)$. This can likely be attributed to the compensating effect of errors in $a_d(\lambda)$ and $a_g(\lambda)$ which leads to reduced errors of $a_{dg}(\lambda)$.

To provide further insight into the performance of the model, we show the *MR* values along with the first and third quartile ratios (QR_1 and QR_3, respectively) of model-derived optimal solutions to the measured values of $a_d(\lambda)$, $a_g(\lambda)$, $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $a_p(\lambda)$ at the four basic wavelengths (Figure 8). The *MR* values (which are equivalent to the second quartile) for the derived-to-measured ratios vary within 1 ± 0.15 for $a_d(\lambda)$, 1 ± 0.2 for $a_g(\lambda)$, and 1 ± 0.1 for $a_{ph}(\lambda)$. With an increase in light wavelength, the *MR* values for $a_d(\lambda)$ show a spectral trend to increase from below the value of 0.9 at 412 nm to ~1.15 at 555 nm (Figure 8a). The *MR* values for $a_g(\lambda)$ show the opposite trend. The spectral pattern of *MR* for $a_{dg}(\lambda)$, $a_{ph}(\lambda)$, and $a_p(\lambda)$ is relatively flat. The QR_1 and QR_3 values for all these absorption coefficients generally follow a similar spectral behavior to the *MR* values.

The results presented in Figures 5–8 demonstrate the capability of the GSCM to perform well with input spectra of $a_{nw}(\lambda)$ with negligible NIR absorption, even though nonzero values of NIR $a_d(750)$, up to 30% of the magnitude of $a_d(440)$, are allowed by the model formalism.

4.2. Evaluation Based on Modified Data of $a_{nw}(\lambda)$ with Nonzero NIR Absorption

In many aquatic environments, particularly in coastal or estuarine waters such as the Chesapeake Bay where particles of terrestrial origin significantly affect the optical properties, the nonalgal particulate absorption

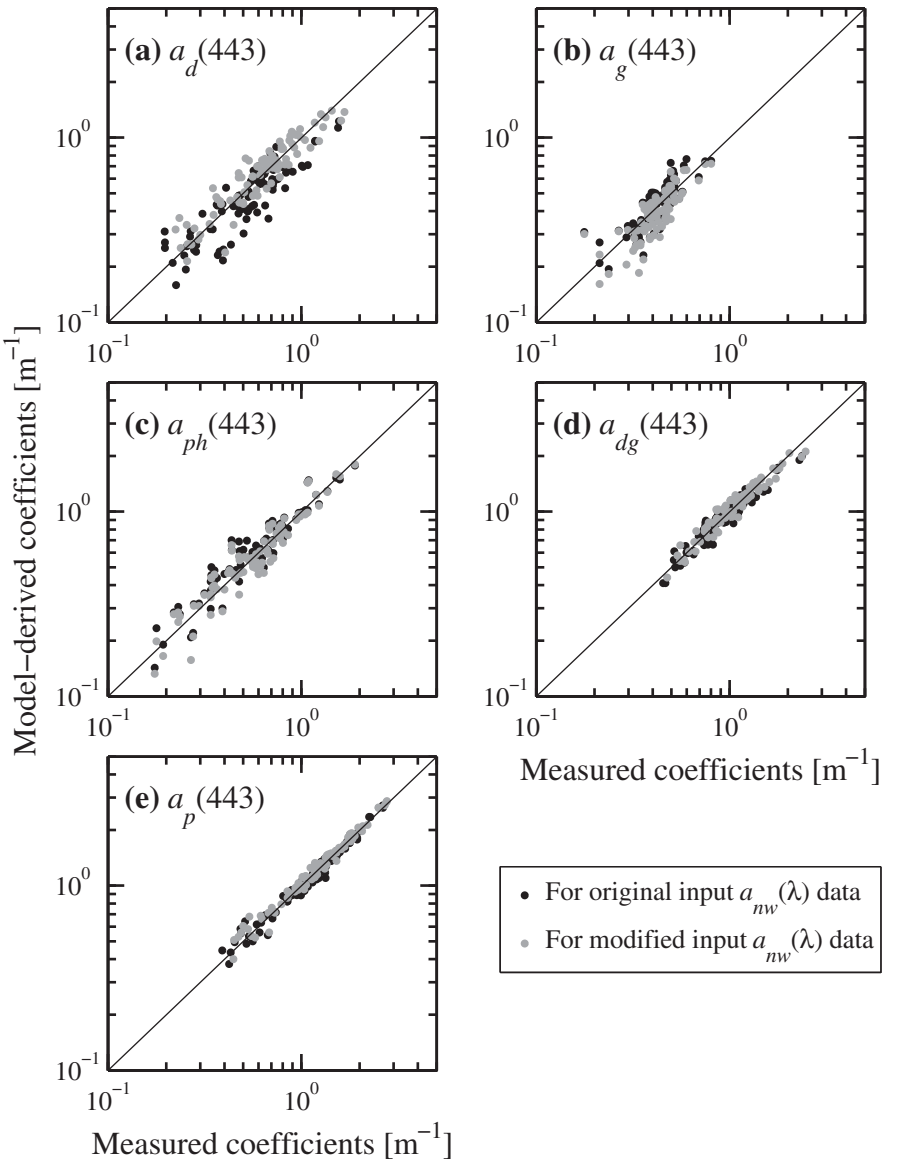

Figure 9. Comparisons of model-derived optimal solutions and measured values of the various absorption coefficients at a light wavelength of 443 nm for the two scenarios of input $a_{nw}(\lambda)$ data used in the model, i.e., original $a_{nw}(\lambda)$ data from our Chesapeake Bay data set (black data points, same data as shown in Figure 6) and modified $a_{nw}(\lambda)$ data (gray data points). The modified data were generated by shifting up each spectrum of nonalgal particulate absorption $a_d(\lambda)$ in the original data set using a random NIR offset. The coefficients of $a_{dg}(\lambda)$, $a_p(\lambda)$, and $a_{nw}(\lambda)$ in the modified data set were recalculated accordingly with the offset-added data of $a_d(\lambda)$ (see text for more details).

coefficient, $a_d(\lambda)$, is likely to exhibit sizable magnitude at NIR wavelengths. This is why our model was designed to include the parameter *B*, which accounts for potential nonzero magnitude of $a_d(\lambda)$ and its changes in the NIR. Unfortunately, we cannot test how effective this parameter is for the performance of the model with the original data from the Chesapeake Bay because of methodological limitations of absorption determinations including the "null-point" in the NIR. Therefore, for such tests we modified the original data set of 90 measurements of absorption coefficients so that the spectra of $a_d(\lambda)$ account for the potential presence of absorption in the NIR.

To construct such a modified data set, a random number, hereafter referred to as N_{rand}, was generated for each $a_d(\lambda)$ spectrum in the original data set. N_{rand} represents possible values of the absorption ratio between the NIR and blue bands, $a_d(750)$:$a_d(440)$, which vary between 0 and 0.3 (see constraint #5 in Table 1). The hypothetical NIR magnitude of absorption represented by the offset parameter *B* was calculated for

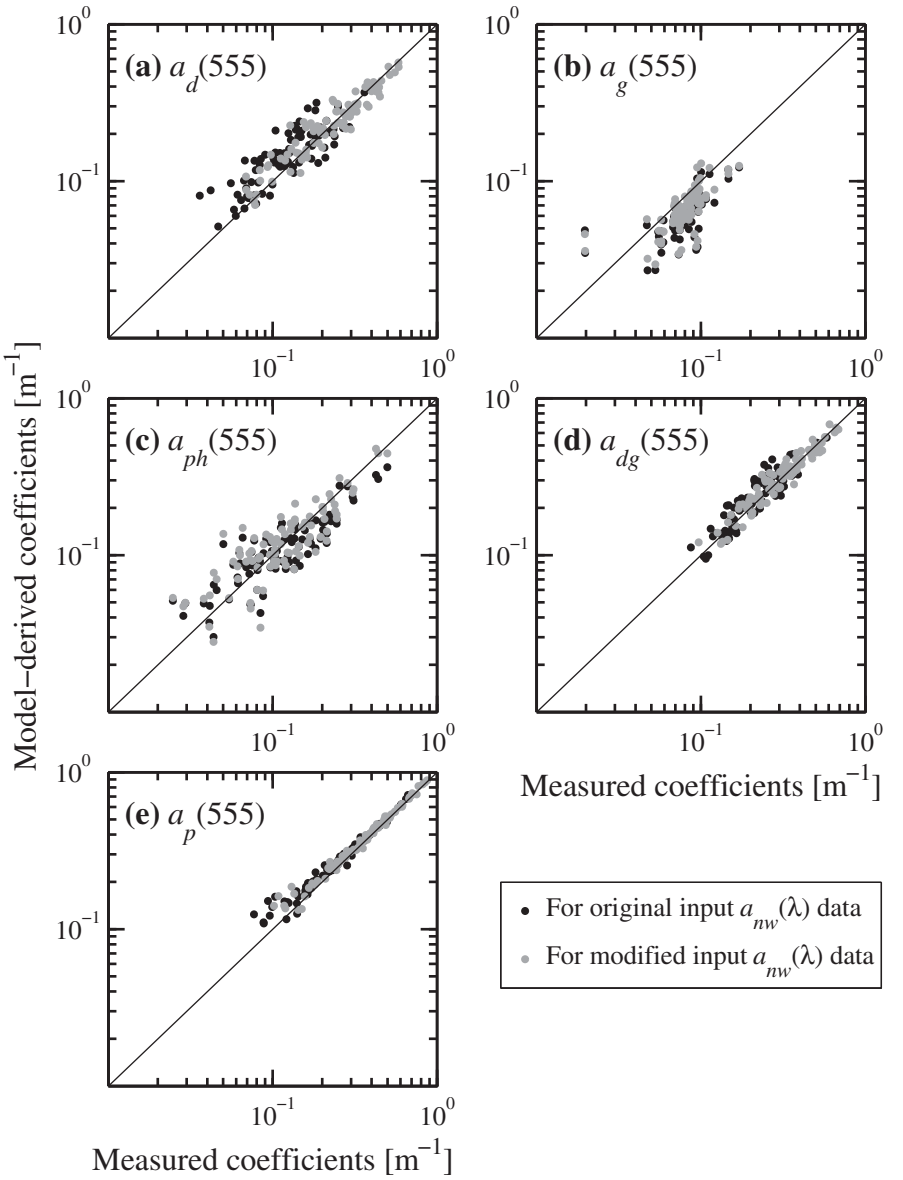

Figure 10. Same as Figure 9 but for the light wavelength of 555 nm.

a given $a_d(\lambda)$ spectrum in such a way that the absorption band ratio between 750 and 440 nm in the offset-added $a_d(\lambda)$ is equal to N_{rand}. The final spectral values of $a_d(\lambda)$ across the entire spectrum were obtained by adding the hypothetical NIR offset to the entire spectrum of the original $a_d(\lambda)$. Data of the $a_{ph}(\lambda)$ and $a_g(\lambda)$ coefficients are independent of $a_d(\lambda)$ and remain unchanged in the modified data set compared with the values in the original data set. Final data for other absorption coefficients containing the nonalgal particulate component, i.e., $a_{dg}(\lambda)$, $a_p(\lambda)$, and $a_{nw}(\lambda)$, were recalculated with the offset-added $a_d(\lambda)$ data. As a result, the input data of spectral $a_{nw}(\lambda)$ are shifted up according to the addition of the random offset. In view of significant presence of mineral particles in the Chesapeake Bay, we expect that the $a_{nw}(\lambda)$ spectra with the added offset might reflect the true absorption properties in the study area better than the original data in which the "null-point" correction was applied.

To minimize the sensitivity of results describing the model performance to the specific set of random numbers N_{rand}, we created a total of five such modified data sets as described above but with different sets of random numbers generated for each data set. We tested our partitioning model GSCM with these

Table 3. Summary of the *MR* and *MPD* Errors for Partitioning Results Obtained With Five Sets of Model Calculations, Each of Which Used a Different Set of Modified Input $a_{nw}(\lambda)$ Data (See Text for More Details)[a]

Variable	*MR* (mean ± std)	*MPD* (%) (mean ± std)
$a_d(412)$	1.047 ± 0.014	11.19 ± 0.79
$a_g(412)$	0.942 ± 0.016	10.73 ± 1.36
$a_{ph}(412)$	0.975 ± 0.011	12.27 ± 0.91
$a_d(443)$	1.040 ± 0.011	11.00 ± 0.97
$a_g(443)$	0.952 ± 0.016	11.24 ± 1.51
$a_{ph}(443)$	0.983 ± 0.006	10.29 ± 0.47
$a_d(490)$	1.064 ± 0.007	13.36 ± 1.29
$a_g(490)$	0.904 ± 0.009	11.86 ± 1.44
$a_{ph}(490)$	0.957 ± 0.005	11.66 ± 0.19
$a_d(555)$	0.992 ± 0.010	11.63 ± 0.78
$a_g(555)$	0.837 ± 0.003	18.77 ± 0.52
$a_{ph}(555)$	1.020 ± 0.014	23.62 ± 1.31

[a]*MR* and *MPD* (see Table 2 for definitions) between model-derived and measured values were first calculated for each of the five sets of model calculations, and the values shown in the table are the mean and standard deviation (std) based on all five sets of calculations.

five data sets and found that five data sets were sufficient to obtain statistically representative results for these purposes. Figures 9 and 10 demonstrate example partitioning results for one of the five modified data sets. For comparison, results from the application of GSCM to the original field data set with no NIR absorption are also shown. Figures 9 and 10 suggest that the model is capable of providing reasonably good results with the input of modified $a_{nw}(\lambda)$ data containing the NIR signal. In the modified data set, the three absorption coefficients containing the random NIR offset, i.e., $a_d(\lambda)$, $a_{dg}(\lambda)$, and $a_p(\lambda)$, are shifted upward compared with the values in the original field data set. The model-derived values of these coefficients are consistent with the upward shift of the data (Figures 9a, 9d, 9e, 10a, 10d, and 10e). Naturally, for the two other coefficients unaffected by the NIR offset, i.e., $a_g(\lambda)$ and $a_{ph}(\lambda)$, the agreement between the model-derived and measured values remains essentially the same for the modified and original data sets (Figures 9b, 9c, 10b, and 10c).

Table 3 provides a summary of the model performance based on results obtained with the five modified data sets. These results are obtained by calculating the *MR* and *MPD* of the model-derived to measured values of $a_d(\lambda)$, $a_g(\lambda)$, and $a_{ph}(\lambda)$ for each modified data set and then calculating the mean and standard deviation of these errors. Table 3 suggests that the performance of GSCM with the modified data sets is actually better than that with the original field data set. For example, at the blue wavelength of 443 nm, the systematic errors calculated from the *MR* values for $a_d(\lambda)$ and $a_{ph}(\lambda)$ are $+4.0 \pm 1.1\%$ and $-1.7 \pm 0.6\%$, respectively, for the modified data sets. In contrast, for the original data set the errors are larger, -8.7% and $+5.6\%$ for $a_d(\lambda)$ and $a_{ph}(\lambda)$, respectively. Slight improvement in the systematic error from $+6.4\%$ for the original data set to $-4.8 \pm 1.6\%$ for the modified data sets is also observed for the model-derived $a_g(\lambda)$. With regard to random errors represented by *MPD*, the performance of the model for the modified data sets is generally comparable with that for the original data set, albeit slight improvement can also be noticed. Based on these results, we expect that our model will be capable to perform well in applications with input data of $a_{nw}(\lambda)$ containing the actual nonzero magnitude in the NIR.

5. Conclusions

In this study, we addressed a challenging problem of developing a model for partitioning the total nonwater absorption coefficient of natural waters $a_{nw}(\lambda)$ (i.e., the total absorption coefficient with pure-water contribution assumed to be a priori known and subtracted) into separate component absorption coefficients of phytoplankton, $a_{ph}(\lambda)$, nonalgal particles, $a_d(\lambda)$, and CDOM, $a_g(\lambda)$. We demonstrated that this objective can be achieved with reasonably good results by formulating a model that does not involve highly restrictive assumptions about the output coefficients of $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$. This is an important advancement compared with the existing models that typically utilize restrictive assumptions about the spectral shapes of component absorption coefficients.

The formalism of our model is based on the stacked-constraints approach which has been recently applied to develop a model for partitioning $a_{nw}(\lambda)$ into $a_{ph}(\lambda)$ and $a_{dg}(\lambda)$ (where $a_{dg}(\lambda) = a_d(\lambda) + a_g(\lambda)$) with only weakly restrictive assumptions about the spectral behavior of $a_{ph}(\lambda)$ and $a_{dg}(\lambda)$ [*Zheng and Stramski*, 2013a]. The underlying idea of the approach is to use a number of inequality constraints that must be satisfied simultaneously by the model outputs of component absorption coefficients, which allows to relax restrictive assumptions and account for variations in these coefficients. Our present model, referred to as the

Generalized Stacked-Constraints Model GSCM, expands this approach to provide additional capability to separate the contributions of $a_d(\lambda)$ and $a_g(\lambda)$. Instead of typical assumption of the exponential spectral shape of $a_d(\lambda)$ and $a_g(\lambda)$ in previous models, GSCM uses a library of realistic spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$ determined from field data collected in the Chesapeake Bay. The use of this library ensures the suitability of the model for various water types encountered in the study area. Another key feature is the use of a parameter that quantifies nonalgal particulate absorption in the NIR spectral region. This parameter is particularly important for waters containing considerable amounts of mineral particles which can exhibit significant absorption even in the NIR.

The required input of GSCM includes the values of $a_{nw}(\lambda)$ at a minimum of four light wavelengths, namely 412, 443, 490, and 555 nm. Such undemanding requirement to run the model enables its broad applicability to data obtained from a variety of remote-sensing and in situ platforms. For example, the model can be applied to $a_{nw}(\lambda)$ data derived from the inversion of radiometric measurements [e.g., *Lee et al*., 2002] with multispectral satellite ocean color sensors such as the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), MEdium Resolution Imaging Spectrometer (MERIS), and the Visible Infrared Imaging Radiometer Suite (VIIRS). The capability of GSCM to partition the satellite-derived total absorption coefficient into $a_{ph}(\lambda)$, $a_d(\lambda)$, and $a_g(\lambda)$ can play a critical role for creating new and advancing existing ocean color data products that can be obtained from information on component absorption coefficients, especially phytoplankton functional types and community size structure, concentration of chlorophyll-*a*, primary productivity, concentration of suspended particulate matter, and organic carbon pools associated with particulate and dissolved components. We note, however, that the application of GSCM in this context will be subject to uncertainties in satellite-derived data of $a_{nw}(\lambda)$, which can arise from different sources, for example, sensor calibration, atmospheric correction, and inversion model for estimating $a_{nw}(\lambda)$ from remote-sensing reflectance. One advantage offered by the GSCM is that this model is likely to provide no feasible solutions rather than solutions with gross error when the input data of $a_{nw}(\lambda)$ are subject to gross error.

Although the present version of GSCM has been developed and validated for applications in a specific coastal region (Chesapeake Bay), we expect that the overall formalism of the model can be generally applicable and extensible to diverse optical water types in different regions. To enable such broader applicability the GSCM will likely require some modifications on the basis of field data of component absorption coefficients for the region of interest, especially alterations or extension of the library of representative spectral shapes of $a_d(\lambda)$ and $a_g(\lambda)$, and inequality constraints and their boundaries. It is also important to note that the data from the Chesapeake Bay used in the present study were frequently characterized by significant differences in the spectral shapes between $a_d(\lambda)$ and $a_g(\lambda)$, and departures of these spectra from the exponential shape, especially $a_d(\lambda)$. These features of $a_d(\lambda)$ and $a_g(\lambda)$ facilitate the separation of these absorption components from one another and are expected to occur commonly in many coastal and inland aquatic environments. Therefore, one can expect that GSCM can be readily adaptable to perform reasonably well in such environments. An important task for the near future research will be also to test the GSCM formalism in other environments, especially open ocean waters where the differences in the spectral shapes between $a_d(\lambda)$ and $a_g(\lambda)$ and the departures from the exponential spectral shape of these coefficients can be smaller than in coastal, estuarine, and inland waters.

Acknowledgments

This work was supported by NOAA's Ocean Remote Sensing (ORS) Program. Partial support for D.S. was provided by NASA Ocean Biology and Biogeochemistry Program (grant NNX15AC55G). We thank all scientists and personnel who contributed to the collection and processing of field data of absorption coefficients used in this study. In particular, we thank L. W. Jr. Harding, C. Hu, and A. Mannino who made the data available through the NASA's SeaWiFS Bio-Optical Archive and Storage System (SeaBASS, http://seabass.gsfc.nasa.gov/). We thank D. G. Bowers and two anonymous reviewers for valuable comments. The contents of this article are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of the NOAA or the U.S. Government.

References

Anderberg, M. R. (1973), *Cluster Analysis for Applications*, 359 pp., Elsevier, N. Y.

Babin, M., and D. Stramski (2004), Variations in the mass-specific absorption coefficient of mineral particles suspended in water, *Limnol. Oceanogr.*, *49*, 756–767.

Babin, M., D. Stramski, G. M. Ferrari, H. Claustre, A. Bricaud, G. Obolensky, and N. Hoepffner (2003), Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe, *J. Geophys. Res.*, *108*(C7), 3211, doi:10.1029/2001JC000882.

Bidigare, R. R., M. E. Ondrusek, J. H. Morrow, and D. A. Kiefer (1990), In vivo absorption properties of algal pigments, *Proc. SPIE*, *1302*, 290–302, doi:10.1117/12.21451.

Bowers, D. G., and C. E. Binding (2006), The optical properties of mineral suspended particles: A review and synthesis, *Estuarine Coastal Shelf Sci.*, *67*, 219–230.

Bricaud, A., and D. Stramski (1990), Spectral absorption coefficients of living phytoplankton and nonalgal biogenous matter: A comparison between the Peru upwelling area and the Sargasso Sea, *Limnol. Oceanogr.*, *35*, 562–582.
Bricaud, A., K. Allali, A. Morel, D. Marie, M. Veldhuis, F. Partensky, and D. Vaulot (1999), Divinyl chlorophyll a-specific absorption coefficients and absorption efficiency factors for *Prochlorococcus marinus*: Kinetics of photoacclimation, *Mar. Ecol. Prog. Ser.*, *188*, 21–32.
Ciotti, A. M., and A. Bricaud (2006), 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. Methods*, *4*, 237–253.
Coble, P. G. (2007), Marine optical biogeochemistry, *Chem. Rev.*, *107*, 402–418.
Falkowski, P. G., R. T. Barber, and V. Smetacek (1998), Biogeochemical controls and feedbacks on ocean primary production, *Science*, *281*, 200–206.
Gallegos, C. L., and P. J. Neale (2002), Partitioning spectral absorption in case 2 waters: Discrimination of dissolved and particulates components, *Appl. Opt.*, *41*, 4220–4233.
Geider, R. J., H. L. MacIntyre, and T. M. Kana (1998), A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients and temperature, *Limnol. Oceanogr.*, *43*, 679–694.
Harding, L. W., Jr., and A. Magnuson (2003), Bio-optical and remote sensing observations in Chesapeake Bay, in *SIMBIOS Project 2003 Annual Report*, edited by G. S. Fargion and C. R. McClain, *NASA Tech. Memo. 212251*, pp. 84–97, NASA Goddard Space Flight Cent., Greenbelt, Md.
Harding, L. W., Jr., A. Magnuson, and M. E. Mallonee (2005), SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the Mid-Atlantic bight, *Estuarine Coastal Shelf Sci.*, *62*, 75–94.
Hu, C., Z. Chen, T. D. Clayton, P. Swarzenski, J. C. Brock, and F. E. Muller-Karger (2004), Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: Initial results from Tampa Bay, Florida, *Remote Sens. Environ.*, *93*, 423–441.
Iturriaga, R., and D. A. Siegel (1989), Microphotometric characterization of phytoplankton and detrital absorption properties in the Sargasso Sea, *Limnol. Oceanogr.*, *34*, 1706–1726.
Kishino, M., M. Takahashi, N. Okami, and S. Ichimura (1985), Estimation of the spectral absorption coefficients of phytoplankton in the sea, *Bull. Mar. Sci.*, *37*, 634–642.
Le, C., C. Hu, J. Cannizzaro, and H. Duan (2013), Long-term distribution patterns of remotely sensed water quality parameters in Chesapeake Bay, *Estuarine Coastal Shelf Sci.*, *128*, 93–103.
Lee, Z., K. L. Carder, and R. A. Arnone (2002), Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters, *Appl. Opt.*, *41*, 5755–5772, doi:10.1364/AO.41.005755.
Lin, J., W. Cao, G. Wang, and S. Hu (2013), Approach for determining the contributions of phytoplankton, colored organic material, and nonalgal particles to the total spectral absorption in marine waters, *Appl. Opt.*, *52*, 4249–4257.
Loiselle, S. A., L. Bracchini, A. M. Dattilo, M. Ricci, A. Tognazzi, A. Cózar, and C. Rossi (2009), The optical characterization of chromophoric dissolved organic matter using wavelength distribution of absorption spectral slopes, *Limnol. Oceanogr*, *54*, 590–597.
Mannino, A., M. E. Russ, and S. B. Hooker (2008), Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the U.S. Middle Atlantic Bight, *J. Geophys. Res.*, *113*, C07051, doi:10.1029/2007JC004493.
Mitchell, B. G. (1990), Algorithm for determining the absorption coefficient of aquatic particulates using the quantitative filter technique (QFT), in *Ocean Optics X Proceedings of SPIE*, edited by R. W. Spinrad, pp. 137–148, SPIE, Bellingham, Wash.
Mitchell, B. G., et al. (2000), Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples, in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*, edited by G. S. Fargion and J. L. Mueller, *NASA/TM-2000–209966*, pp. 125–153, NASA Goddard Space Flight Cent., Greenbelt, Md.
Mitchell, B. G., M. Kahru, J. Wieland, and M. Stramska (2002), Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples, in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 3*, edited by J. L. Mueller and G. S. Fargion, pp. 231–257, NASA Goddard Space Cent., Greenbelt, Md.
Mitchell, B. G., M. Kahru, J. Wieland, and M. Stramska (2003), Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples, in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*, edited by G. S. Fargion, J. L. Mueller, and C. R. McClain, *NASA/TM-2003–211621/Rev4*, vol. IV, pp. 39–64, NASA Goddard Space Flight Cent., Greenbelt, Md.
Moore, C., et al. (2004), Detritus, trophic dynamics and biodiversity, *Ecol. Lett.*, *7*, 584–600.
Nelson, J. R., and C. Y. Robertson (1993), Detrital spectral absorption: Laboratory studies of visible light effects on phytodetritus absorption, bacterial spectral signal, and comparison to field measurements, *J. Mar. Res.*, *51*, 181–207.
Roesler, C. S., M. J. Perry, and K. L. Carder (1989), Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters, *Limnol. Oceanogr.*, *34*, 1510–1523.
Röttgers, R., C. Dupouy, B. B. Taylor, A. Bracher, and S. B. Woźniak (2014), Mass-specific light absorption coefficients of natural aquatic particles in the near-infrared spectral region, *Limnol. Oceanogr.*, *59*, 1449–1460.
Schaeffer, B. A., J. D. Hagy, R. N. Conmy, J. C. Lehrter, and R. P. Stumpf (2012), An approach to developing numeric water quality criteria for coastal waters using the SeaWiFS satellite data record, *Environ. Sci. Technol.*, *46*(2), 916–922.
Schofield, O., T. Bergmann, M. J. Oliver, A. Irwin, G. Kirkpatrick, W. P. Bissett, M. A. Moline, and C. Orrico (2004), Inversion of spectral absorption in the optically complex coastal waters of the Mid-Atlantic Bight, *J. Geophys. Res.*, *109*, C12S04, doi:10.1029/2003JC002071.
Stramski, D., A. Bricaud, and A. Morel (2001), Modeling the inherent optical properties of the ocean based on the detailed composition of planktonic community, *Appl. Opt.*, *40*, 2929–2945.
Stramski, D., S. B. Woźniak, and P. J. Flatau (2004), Optical properties of Asian mineral dust suspended in seawater, *Limnol. Oceanogr.*, *49*, 749–755.
Stramski, D., M. Babin, and S. B. Woźniak (2007), Variations in the optical properties of terrigenous mineral-rich particulate matter suspended in seawater, *Limnol. Oceanogr.*, *52*, 2418–2433.
Tassan, S., and G. M. Ferrari (2003), Variability of light absorption by aquatic particles in the near-infrared spectral region, *Appl. Opt.*, *42*, 4802–4810.
Tzortziou, M., J. Herman, C. Gallegos, P. Neale, A. Subramaniam, L. Harding, and Z. Ahmad (2006), Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure, *Estuarine Coastal Shelf Sci.*, *68*, 348–362.
Zheng, G., and D. Stramski (2013a), A model based on stacked-constraints approach for partitioning the light absorption coefficient of seawater into phytoplankton and non-phytoplankton components, *J. Geophys. Res. Oceans*, *118*, 2155–2174, doi:10.1002/jgrc.20115.
Zheng, G., and D. Stramski (2013b), A model for partitioning the light absorption coefficient of suspended marine particles into phytoplankton and nonalgal components, *J. Geophys. Res. Oceans*, *118*, 2977–2991, doi:10.1002/jgrc.20206.